1966

Optimality considerations of the graphical evaluation and review technique (gert)

Gary Werberger
Lehigh University

OPTIMALITY CONSIDERATIONS OF THE GRAPHICAL EVALUATION AND REVIEW TECHNIQUE (GERT)

by

Gary Werberger

A Thesis

Presented to the Graduate Faculty

Of Lehigh University

In Candidacy for the Degree of

Master of Science

Lehigh University

1966

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

May 5, 1966
Date

Gary E. Whitehouse
Professor in Charge

Arthur F. Gould
Head of the Department

ACKNOWLEDGEMENTS

My sincere appreciation is expressed to the members of the Industrial Engineering faculty of Lehigh University headed by Professor A. F. Gould, and to the many members of the Western Electric Engineering Research Center staff whose cooperation and assistance formed a major contribution to this thesis.

In particular, the guidance and encouragement of Professor G. E. Whitehouse, who was, in part, originally responsible for the development of GERT, is acknowledged. As members of my advisory committee, Professors S. Monro and W. T. Richardson offered helpful suggestions. Mr. E. S. Bahary of the Engineering Research Center staff deserves special recognition for his cooperation and support. The discussions with Mr. V. J. Graziano of the Western Electric Finance Division were also very beneficial.

This thesis could not have been completed without the continued encouragement of my wife, Elinor.

ABSTRACT

Graphical Evaluation and Review Technique (GERT) is a recently developed tool that combines the disciplines of flowgraph theory, moment generating functions, and PERT to give graphical solutions in many problem areas. The major contribution of this paper is to extend the application of GERT in the area of optimality analysis.

The intent of this work is to provide an analytic structure for a decision-making system that is more general, more descriptive, and more directly computational than has heretofore been possible without the use of GERT. It is based on the Markov process as a system model, that contains both probabilistic and decision-making features, and utilizes GERT to reinforce the iterative technique of R. A. Howard (7), that is similar to dynamic programming as an optimization method.

The sensitivity of this iterative technique is investigated, as the Markov structure is varied. Other, less sensitive, iterative procedures are analyzed and discussed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (continued)

CHAPTER I

INTRODUCTION

An increase in the use of Network analysis has been due largely to the ease with which systems can be modeled in network form. This, coupled with the need for a communication mechanism to discuss, analyze and schedule an operating system, has led to the introduction of a new graphical problem solving technique which is called GERT: Graphical Evaluation and Review Technique. GERT is a procedure which combines the disciplines of flow graph theory, moment generating functions and PERT to obtain solutions to stochastic problems. It has been shown by Pritsker, Happ and Whitehouse (17,18,23), that complex systems and problems can be analyzed by this procedure in a more direct manner than ever before possible.

The importance of such a procedure is therefore obvious. It now becomes increasingly more important to consider new and diverse areas of application for GERT. This thesis is concerned primarily with extending the application of GERT to optimal decision making. GERT, in its present form, is a reporting technique. It is desirable to utilize the concepts of GERT and redefine an iterative procedure to obtain optimal results when alternative decisions are inherent in an operational system under study.

Chapter II is concerned with defining the terminology and relationships of traditional Markov Processes and problems that are solved using GERT. It will be shown that the quantities required in the solution of a Markov Process are easily solved for using the method employed by GERT. Howard (7) has extended Markov Theory by introducing an iterative technique that determines the alternatives which optimize the Process. Since Markov Theory and GERT are closely related, the iterative technique developed by Howard becomes a basis on which an approach is built to utilize GERT as an optimizing as well as a reporting method.

The concepts of reward, reward structure and alternative policies are introduced in Chapter III and Chapter IV with the aid of the Toymaker Problem and the Taxicab Operation as examples from Howard (7, pp 18 and 44). The reward structure and alternative policies are necessary quantities in defining problems of optimization. The rate of return (or gain) of a system, and expected immediate reward of a particular state are defined and solved by graphical techniques. These parameters are then used in Howard's policy-iteration routine to solve the above examples.

The real power of GERT is demonstrated in Chapter V where it is shown that problems of a more general nature, such as continuous and semi-Markov processes, and be solved in the same manner as the simpler, more straight forward, discrete-time Markov process. The policy-iteration technique is used similarly to find the optimal solution to these more general types of problems. This is demonstrated by an analysis of a modification to the toymaker problem. The sensitivity of the policy iteration technique is investigated by varying the reward structure of the example.

In the sixth and last chapter, conclusions of this paper are enumerated, as well as suggested areas for future study.

CHAPTER II

RELATIONSHIP OF GERT TO FINITE MARKOV PROCESSES

An introductory review of the concepts and terminology of Markov Theory is necessary to provide a common ground for further discussion. Kemeny and Snell (11) give a complete and readable treatment of the material and this review is based on their text.

A finite Markov chain is a stochastic process which moves through a finite number of states, and for which the probability of entering a certain state depends only on the last state occupied. It starts in state S_j with probability $p_j(o)$. If at any time it is in state S_i, then it moves on the next step to S_j with probability p_{ij}. The initial probabilities are thought of as giving the probabilities for the various possible starting states. The initial probability vector $\Pi_o = p_j(o)$ and the transition matrix $P = [p_{ij}]$ completely determine the Markov chain process and its stochastic nature.

The states in a Markov process can be considered to be divided into two classifications; transient and ergodic sets. The former, once left are never again entered, while the latter are never left after once being entered. If there is only one element in an ergodic set, it is called an absorbing state. States are therefore classified as a transient state if it is included in a transient set, and an ergodic state if it is included in an ergodic set. An ergodic state is further subclassified as an absorbing state if there is only one state in an ergodic set.

This classification of states is used to classify chains. In an absorbing chain, all non-transient states are absorbing. Figure 1 gives the transition matrix and corresponding GERT network of such a chain. States S_3-S_6 are transient states and states S_1 and S_2 are absorbing states.

A regular chain is one that has no transient sets and has a single ergodic set. Thus, no matter where the process starts, after a sufficient lapse of time, it could be in any state. The transition matrix and corresponding GERT network of a regular chain is shown in Figure 2. From the graphical description, it becomes clear that it is possible to be in any state after a given number of transitions.

A regular chain is only one specific type of the more general ergodic chain where the chain is composed of d cyclic sets. For a given starting position it will move through the cyclic sets in a definite order, returning to the set of the starting state after d steps. Turning now to Figure 3, it becomes clear that starting from an even-numbered state, the process can be in even-numbered states only in an even number of steps, and in an odd-numbered state in an odd number of steps; hence the even and odd states form two cyclic classes.

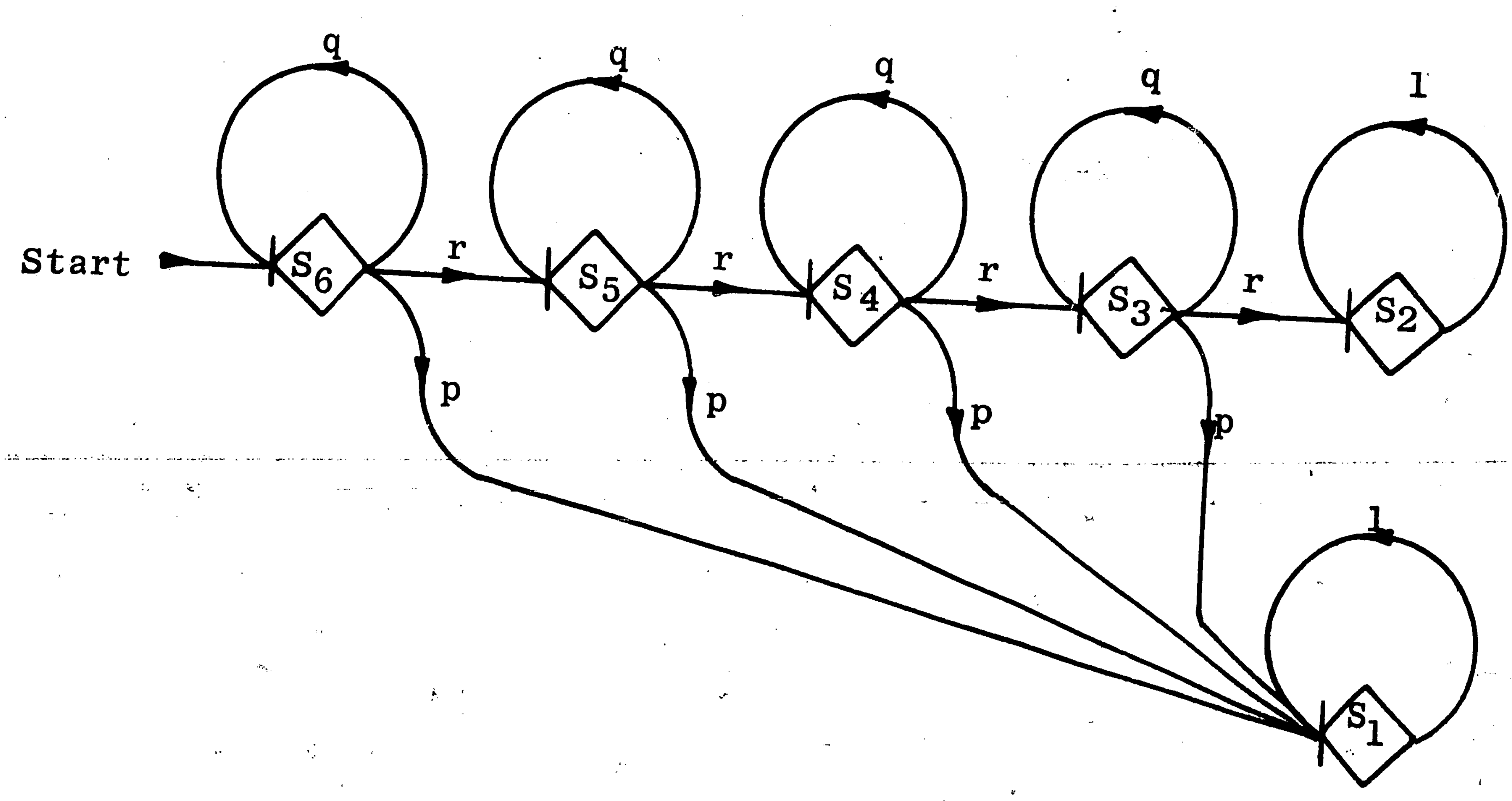
From the above discussion, it is noted that every finite Markov chain must have an ergodic set, but there need be no transient set. If a chain has more than one ergodic set, then there is absolutely no interaction between these sets. These chains may therefore be

COLLEGE MATRICULATION PROBLEM

A student going to a certain college has each year a probability p of flunking out, a probability q of having to repeat the year, and a probability r of passing to the next year.

$$
P = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \end{array}
\begin{array}{c}
\begin{array}{cccccc} S_1 & S_2 & S_3 & S_4 & S_5 & S_6 \end{array} \\
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
p & r & q & 0 & 0 & 0 \\
p & 0 & r & q & 0 & 0 \\
p & 0 & 0 & r & q & 0 \\
p & 0 & 0 & 0 & r & q
\end{bmatrix}
\end{array}
$$

Transition Matrix



GERT Network

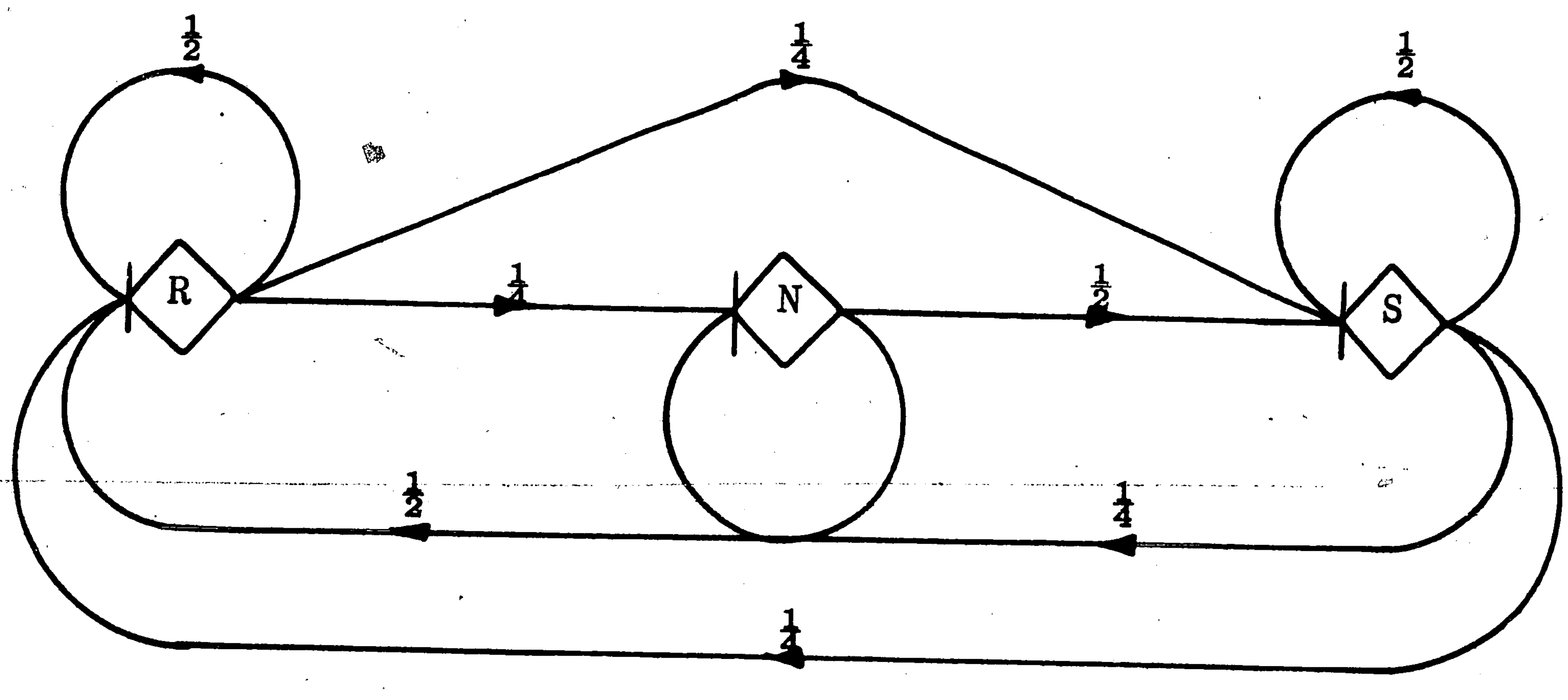
Figure 1. Absorbing Chain

LAND OF OZ PROBLEM

In the land of Oz they never have two nice days in a row. If they have a nice day, they are as likely to have snow as rain the next day. If they have snow (or rain) they have an even chance of having the same the next day. If there is a change from snow or rain, only half of the time is this a change to a nice day.

$$P = \begin{array}{cc} & \begin{array}{ccc} R & N & S \end{array} \\ \begin{array}{c} R \\ N \\ S \end{array} & \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \end{array}$$

Transition Matrix



GERT Network

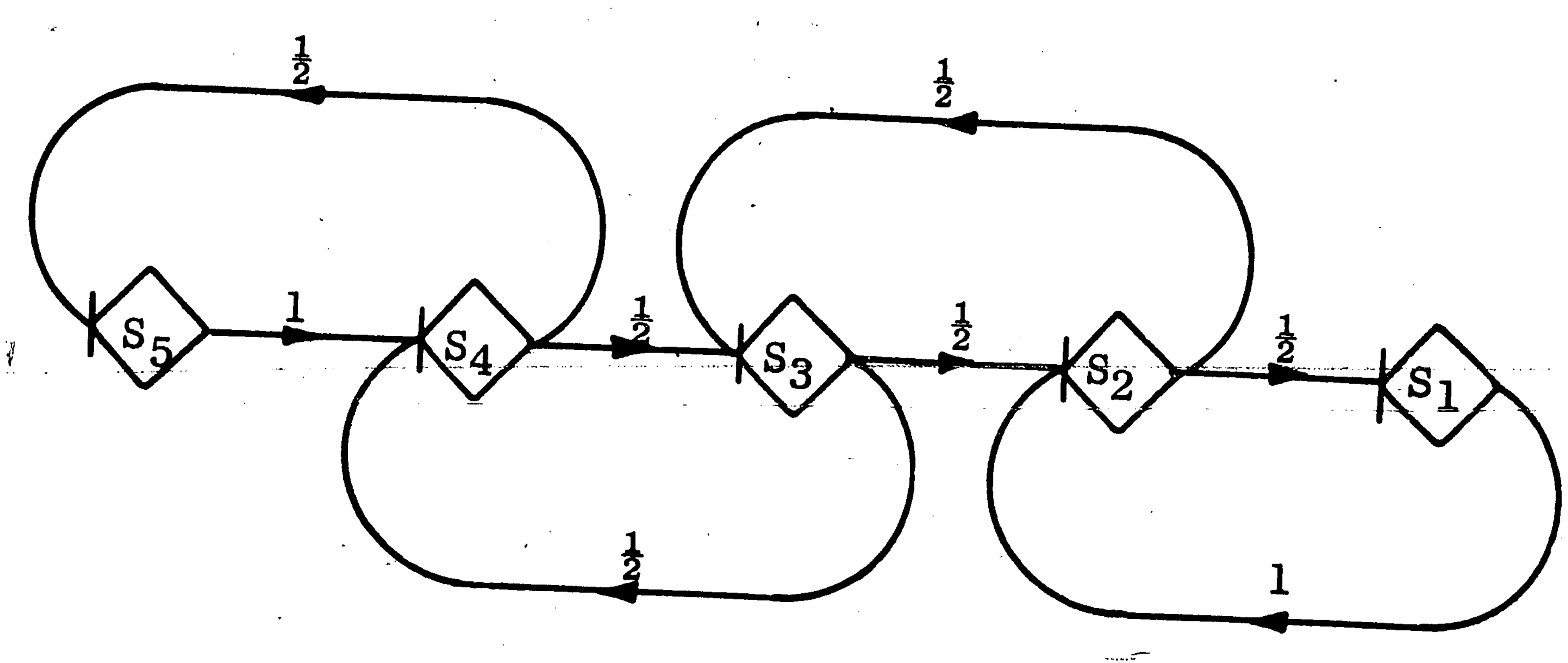
Figure 2. Regular Chain

RANDOM WALK PROBLEM

The random walk problem has from 0 to n states, and states 0 and n are non-absorbing. the problem can be thought of as two men playing a game of chance in which player A has the probability p of winning - assume one dollar is bet each time the game is played. When one player has won all the money, one dollar is given back to the losing player so that the game might continue.

$$
P = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array}
\begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}
$$

Transition Matrix



GERT Network

Figure 3. Ergodic Chain

studied separately, and hence, without loss of generality, the entire chain is assumed to be a single ergodic set. In chains containing transient states, the process always moves toward the ergodic sets. The probability that the process is in an ergodic set tends to 1 after a large number of transitions; and it cannot escape from an ergodic set once it enters it. Here again, it is emphasized that the ergodic set is of primary importance in solving the many questions desired from the system.

The questions one is normally interested in answering are:

(1) The probability of being in a particular state after n transitions, starting from a predetermined state.

(2) The change in the probability of being in a particular state after n transitions when the starting state is changed.

(3) The mean and variance of the number of steps (or transition time) necessary for the process to go from state S_i to S_j. The probability that the process passes through state S_k.

(4) An observation of the process only when it is in a certain subset of states.

The analytical techniques used in solving the above questions are well known, especially when the transition time between states is a discrete function such as the college matriculation problem where the transition time between states is one year. The reader is referred to Kemeny and Snell (11) and Takacs (22) for the development and application of these techniques. The continuous

Markov process is similar to the discrete-time Markov process. This more generalized description of a Markov process is defined as a process in which the transition time between states is a random variable taken from the exponential distribution with a known mean. The methodology used in solving the continuous time Markov process can be found in Howard (7).

In recent literature, there has been much written about a closely related probabilistic system. At the International Congress held in Amsterdam in 1954, Levy (12) and Smith (21) independently presented papers in which a new class of stochastic processes, called Semi-Markov processes, was defined. The process was a generalization of both continuous and discrete parameter Markov processes. Analytical techniques devised to analyze this more general class of processes, however, are extremely complex.

<u>An Introduction to GERT with "Exclusive-or" and "Branch" Nodes</u>

The purpose of this section is to review the work of Pritsker, Happ and Whitehouse (17,18,23) in their development of "Graphical Evaluation and Review Technique" (GERT). GERT and the relationship to techniques used in analyzing Markov chains is emphasized.

GERT is a multiparameter network which allows random variables to be placed on the branches of the network. These random variables, referred to as transmittances, are multi-dimensional and are composed of the probability of taking a given path, p_{ij}, and the distribution of the time to traverse that path, $f_{ij}(t)$. The transmittance in a GERT network is always a directed path. The sum of all the conditional probabilities of the branches emanating from a node must equal one. t is allowed to be a random variable. However, in a discrete-time Markov process, $f_{ij}(t)$ is a constant. The transition probabilities and time distribution are the same quantities that describe a Markov process.

The GERT network also allows logical nodes to be placed in the graph. Of all the types of logical nodes that are discussed by Whitehouse (23, pp. 23-24), the "exclusive-or" and "branch" nodes are the most descriptive of the states in a Markov process. The "exclusive-or" node is shown in Figure 4. 3 is described as occurring if a occurs but not b, and vice versa. The "branch" node is also illustrated in Figure 4. This is a node at which the system may transfer along one path or the other with known probabilities. Since these types of nodes are descriptive of the Markov process, the review of GERT is restricted to applications in networks containing only "exclusive-or" and "branch" nodes.

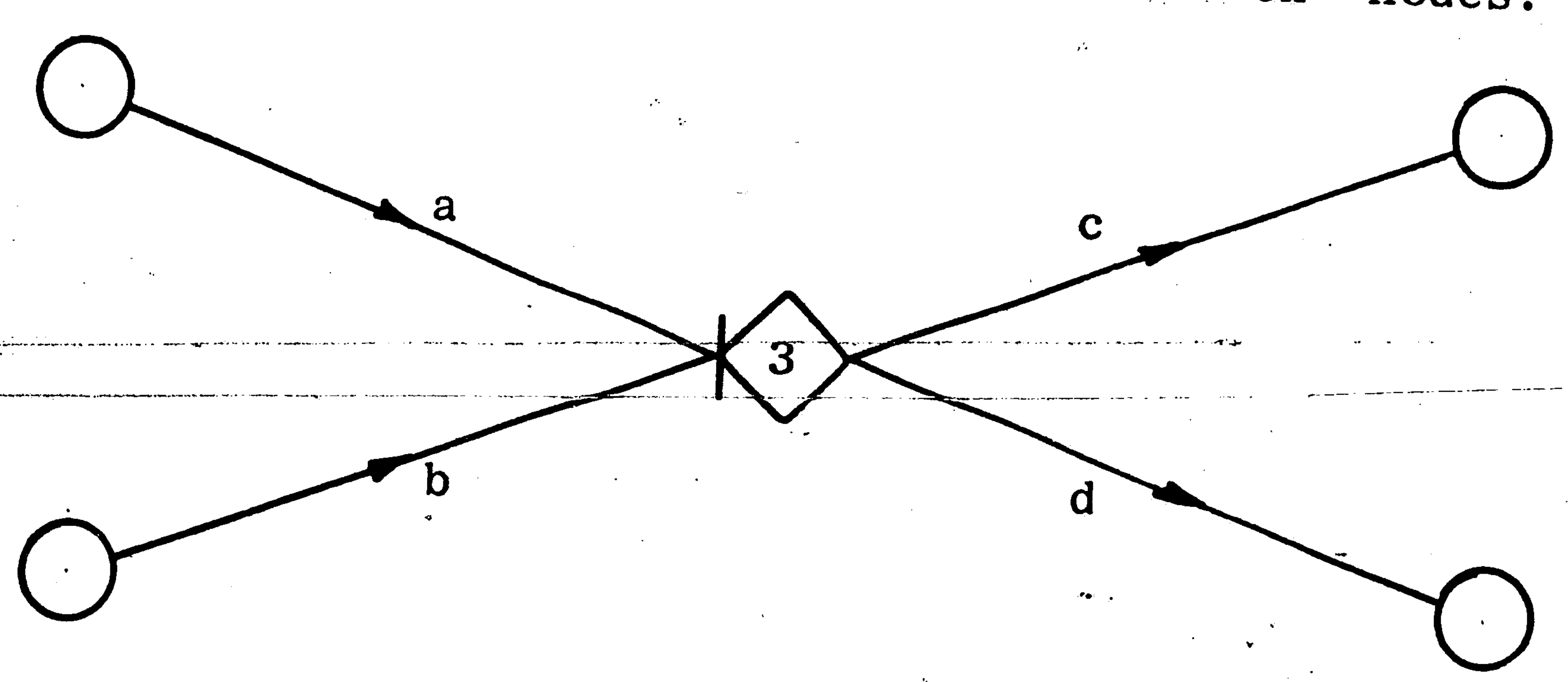


Figure 4. "Exclusive-or" and "Branch" Node

Since the three basic elements - elements in series, elements in parallel, and loops - can be made to behave in a GERT network as they do for the flowgraph, the powerful topological equation can be employed to reduce the GERT network as it is used to reduce flowgraphs. In general, the topological equation takes this form: The sum of the loops formed in the following manner will always equal zero in a closed flowgraph.

$$\sum \text{loops} = 1 - L_1 + L_2 - L_3 + \ldots + (-1)^i L_i + \ldots = 0$$

where L_i is the sum of the i^{th} order loops. A first order loop is a consecutive path of arrows leading from a node and returning to the same node. The value of such a loop is the product of the transmittances around that loop. There are n-order loops which can be described as the product of n non-touching first order loops. The topological equation is defined for closed flowgraphs which are composed entirely of loops. The topological relationship is demonstrated for the flowgraph shown in Figure 5 as follows:

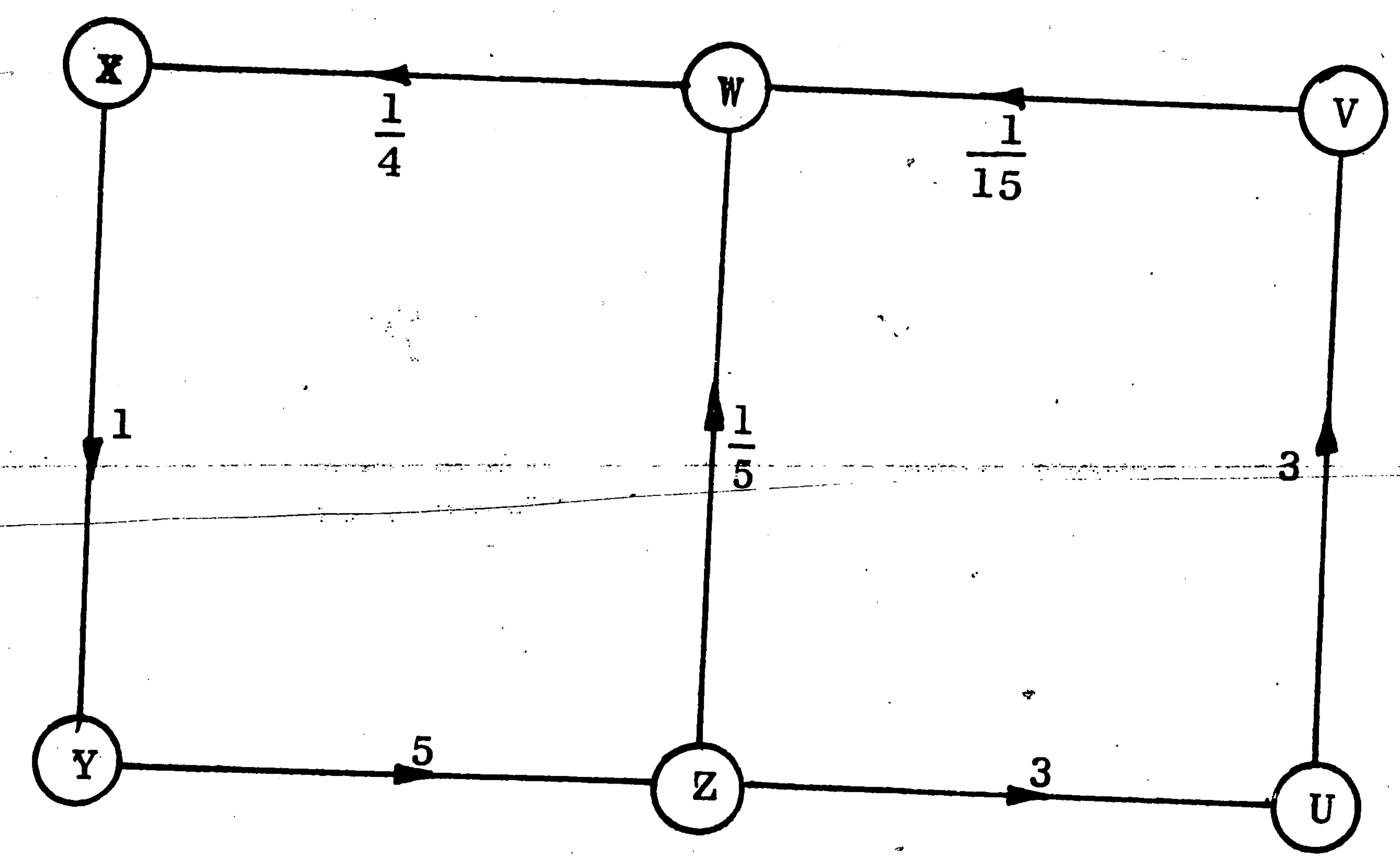


Figure 5. Demonstration of the Topological Relationship In Flowgraphs

$$\text{Loop I (XYZW)} = 1 \times 5 \times \frac{1}{5} \times \frac{1}{4} = \frac{1}{4}$$

$$\text{Loop I (XYZUVW)} = 1 \times 5 \times 3 \times 3 \times \frac{1}{15} \times \frac{1}{4} = \frac{3}{4}$$

$$L_1 = \frac{1}{4} + \frac{3}{4} = 1$$

and $\sum \text{Loops} = 1 - L_1 = 0$

The restriction that the topological equation holds only for closed flowgraphs does not, however, prevent its use on open flowgraphs because the graph can be closed with a fictitious transmittance. The flowgraph can be solved for the transmittance T between the two nodes that have been artificially closed. An approach in finding the equivalent transmittance T between two nodes of interest is using Mason's Rule:

$$T = \frac{(\sum(\text{path} \times \sum \text{non touching loops}))}{\sum \text{loops}}$$

This equation is stated as follows: Write down the product of transmittances along each path from the one node of interest to the other. Multiply its transmittance by the sum of the non-touching loops. Sum these modified path transmittances and divide by the sum of all the loops in the open flowgraph.

The flowgraph is a multiplicative system. Since the probability elements of the GERT network are multiplicative, they conform to the flowgraph system. The time variable is additive in nature. Therefore, the moment generating function was chosen to transform the time variable into a multiplicative parameter. The total transmittance between two nodes i and j in a GERT network is the product of the probability of realizing node j given that node i is occupied and the MGF of the time function to realize that node.

$$W_{ij}(s) = p_{ij}M_{ij}(s)$$

Utilizing GERT to Solve Markov Chain Problems

Returning to the original ideas of absorbing chains and ergodic chains, it is interesting to note that they resemble open and closed flowgraphs. Further, a state can be thought of as an "exclusive-or" and "branch" type node, and the transition from state to state defined by the transmittance.

GERT by nature of its development, represents processes which can be assumed to terminate in some node, which in Markov terminology could be described as an absorbing state. It therefore seems reasonable to assume that GERT is ideally suited to the investigation of processes with absorbing chains. Although ergodic chains cannot be represented by a typical GERT Network, much information can be obtained by forcing it into the familiar GERT format. The network of the ergodic Markov process is forced into GERT format by breaking one node into two nodes. One node acts as the source of the GERT network and has only activities emanating from it. The other has only nodes entering it. This node will be the sink of the GERT network.

The questions posed earlier on absorbing and ergodic Markov chains can be answered by adapting the problem to a GERT network and employing topology equations. Examples of this procedure can be found in Whitehouse (23, Chapter 7).

Once the transmittance, $W_e(s)$, is determined between two states of interest, the probability, p_e, of realizing the portion of the graph being investigated is given by the relationship:

$$p_e = W_e(s) \Big]_{s = 0}$$

since all MGF become equal to one when s is equated to zero. The MGF will be the quotient of the output of the graph divided by p_e:

$$M_e(s) = \frac{W_e(s)}{p_e}$$

The first derivative of $M_e(s)$ evaluated at s equal to zero will yield the mean of the transition time between the two states of interest, and the second derivative of $M_e(s)$ evaluated at s equal to zero will yield the variance. In general, the n^{th} derivative will yield the expected value of the n^{th} power of t.

Some of the inferences that can be made concerning the system under study utilizing the above information will become evident in the following chapters. For example, the Toymaker problem, is analyzed in detail in the following chapter utilizing GERT representation. The problem is modified according to Howard (7) and structured as an optimization problem. It is then shown that GERT can be used directly in solving for the parameters necessary in Howard's approach for finding the optimal solution.

CHAPTER III

THE TOYMAKER PROBLEM

The toymaker problem is a very simple example of a discrete-time Markov process. The toymaker is involved in the novelty toy business. His situation is described as follows:

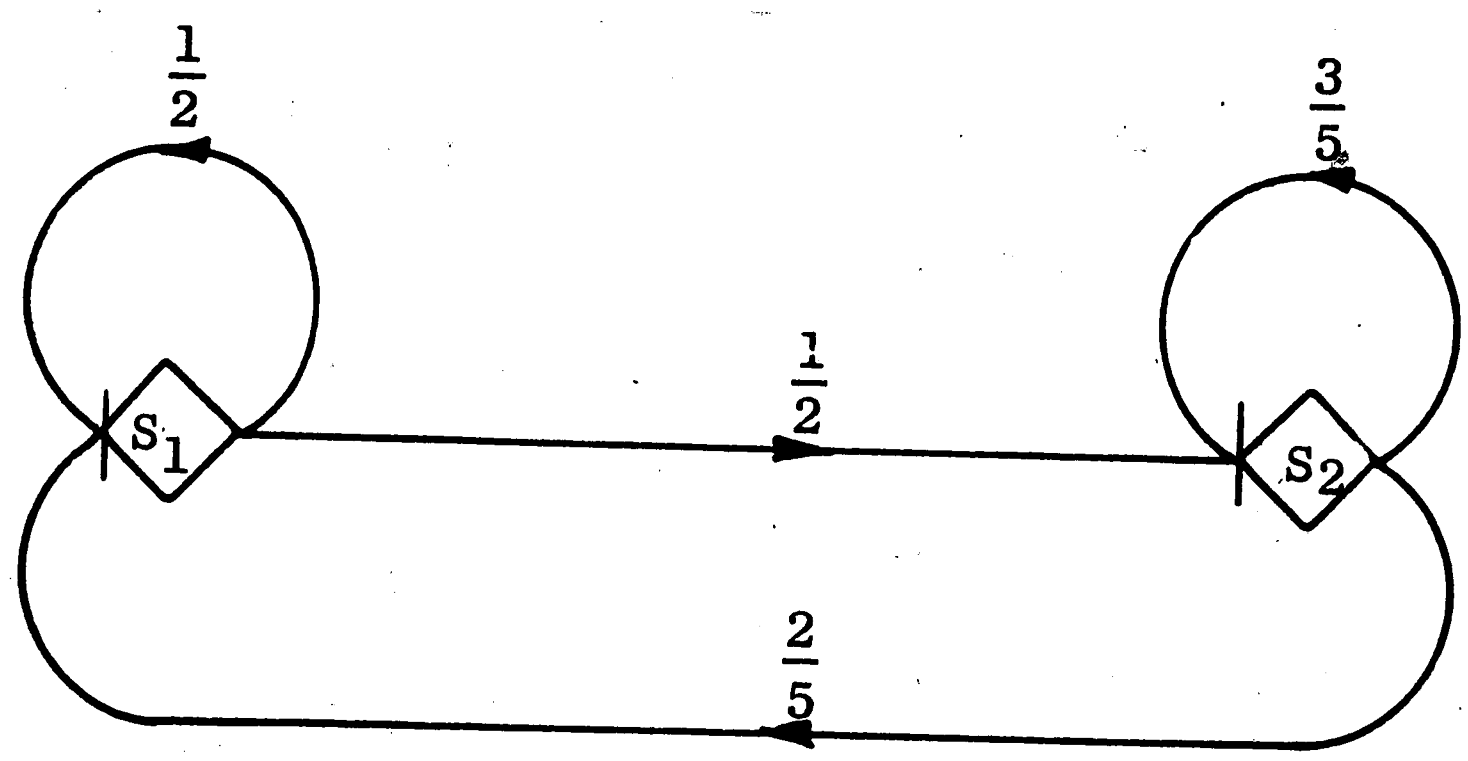
> The toymaker may be in either of two states. He is in the first state if the toy he is currently producing has found great favor with the public. He is in the second state if his toy is out of favor. When he is in state 1, there is 50% chance of his remaining in state 1 at the end of the following week and, consequently, a 50% chance of an unfortunate transition to state 2. When he is in state 2, he experiments with new toys, and he may return to state 1 with probability 2/5 or remain unprofitable in state 2 with probability 3/5.

The transition matrix is shown in Figure 6. This matrix is called a stochastic matrix, since it exhibits the properties of the rows summing to 1, and the elements are non-negative and not greater than 1 ($0 \leq p_{ij} \leq 1$).

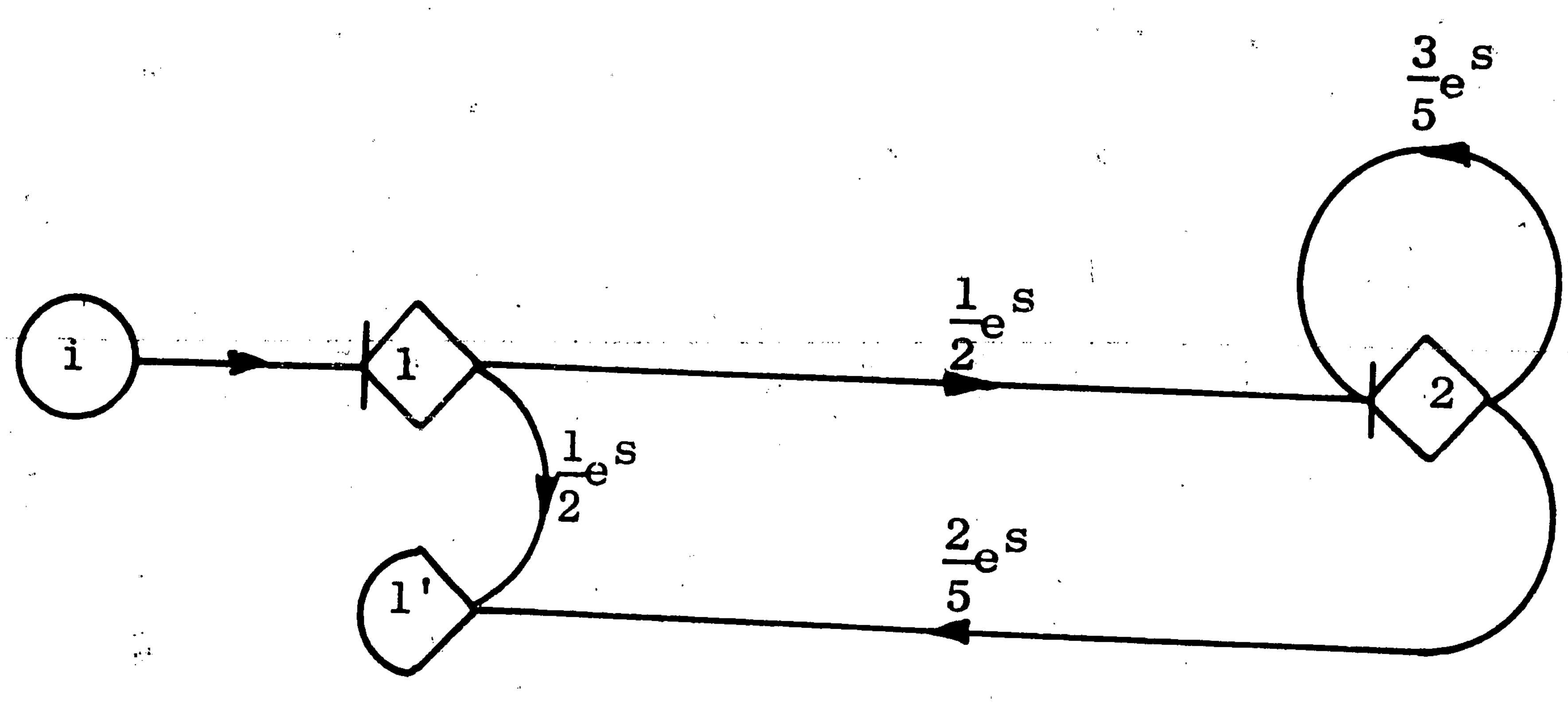
Traditional Markov theory studies the transient behavior of the discrete-time Markov process by describing the time function in terms of the Z-transform for theoretical convenience. The Z-transform is useful in Markov processes because the probability transients in Markov processes are geometric sequences. The Z-transform provides a closed form expression for such sequences. This technique enables one to solve for the limiting state probabilities; that is, the probability of being in a particular state after a large number of transitions. An ergodic process such as the toymaker problem exhibits the property that the state occupancy probabilities are

$$P = \begin{array}{c c} & \begin{array}{cc} S_1 & S_2 \end{array} \\ \begin{array}{c} S_1 \\ \\ S_2 \end{array} & \left[\begin{array}{cc} \frac{1}{2} & \frac{1}{2} \\ \\ \frac{2}{5} & \frac{3}{5} \end{array} \right] \end{array}$$

Transition Matrix



Network Form



GERT Representation to Calculate Mean Recurrence Time

Figure 6. The Toymaker Problem

independent of the starting state of the system if the number of state transitions is large.

Whitehouse (23, pp. 132) has suggested that networks of ergodic Markov chains such as the toymaker problem can be forced into GERT format by breaking one node into two nodes. One node acts as the source of the GERT network and only has transitions emanating from it. The other has transitions entering it, and acts as the sink of the GERT network. This reduces the network to a finite problem. According to Whitehouse, the approach is to find the equivalent transmittance between the split nodes. The Mean Recurrence Time, μ_{rt}, is obtained by taking the first derivative of the resulting MGF and evaluating it at s equal to 0.

The analysis of the toymaker problem is begun by describing the parameters that comprise the transmittances of the network. These transmittances are shown graphically on the GERT representation in Figure 6 and were calculated as follows:

$$W_{ij}(s) = p_{ij}M_{ij}(s)$$

where $M_{ij}(s) = e^{ts}$

and t is the discrete time interval from state i to state j.

The discrete time interval from state i to state j in the toymaker problem is one week.

Referring to the GERT representation of the problem in Figure 6, the equivalent transmittance between state 1 and state 1' is solved by the topological relationship:

$$W_e(s) = \frac{\frac{1}{2}e^s\left[1 - \frac{3}{5}e^s\right] + \frac{1}{5}e^{2s}}{\left[1 - \frac{3}{5}e^s\right]} = \frac{\frac{1}{2}e^s - \frac{1}{10}e^{2s}}{\left[1 - \frac{3}{5}e^s\right]}$$

$$p_e = W_e(s)\Big|_{s=0} = \left.\frac{\frac{1}{2}e^s - \frac{1}{10}e^{2s}}{\left[1 - \frac{3}{5}e^s\right]}\right|_{s=0} = 1$$

$$M_e(s) = \frac{W_e(s)}{p_e} = \frac{\frac{1}{2}e^s - \frac{1}{10}e^{2s}}{\left[1 - \frac{3}{5}e^s\right]}$$

The mean recurrence time is calculated to be:

$$\mu_{rt} = \left[\frac{d}{ds} M_e(s)\right]_{s=0} = \frac{9}{4}$$

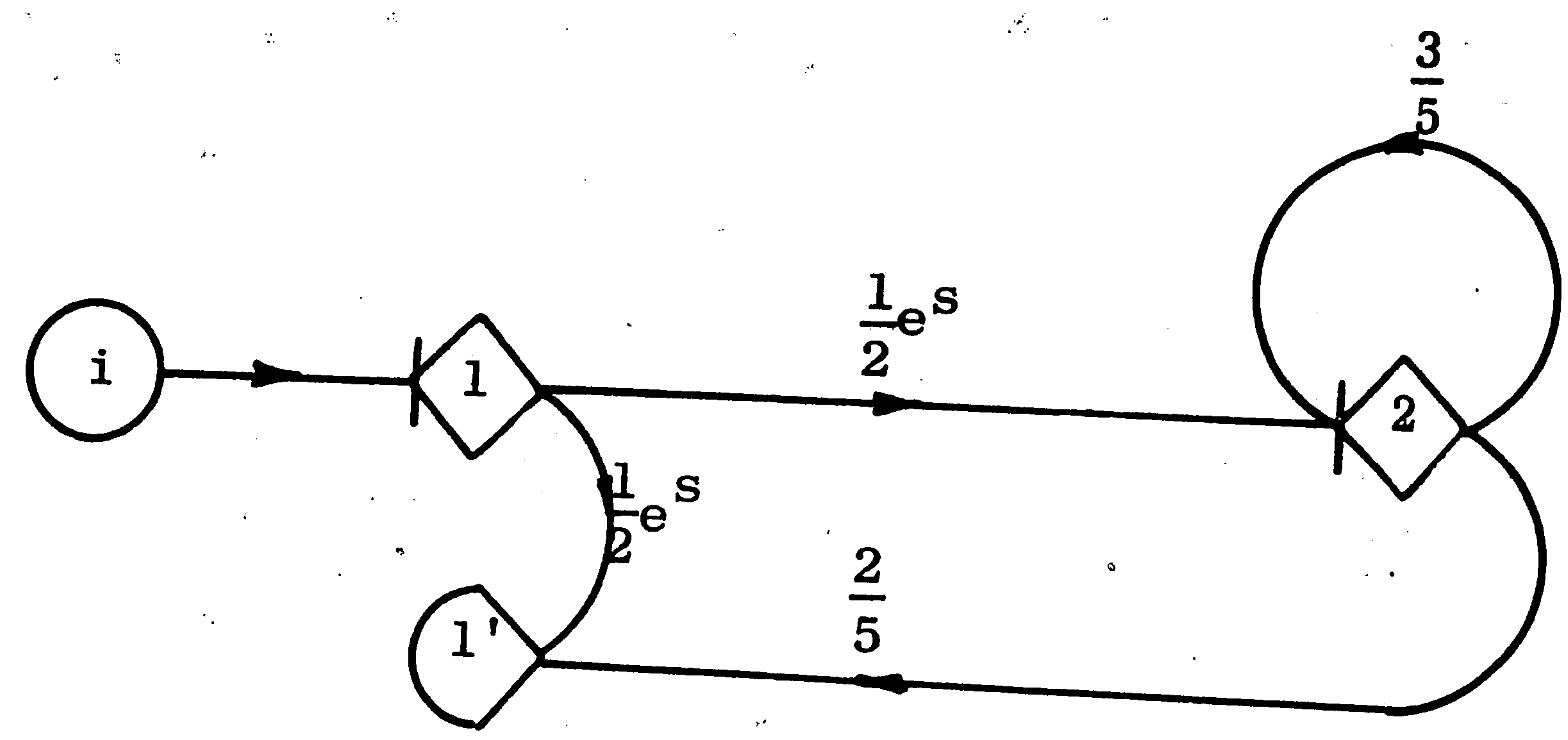
To determine the stationary probability of being in a given state or node, the MGF in all paths except those leaving the node of interest are equated to one. The GERT representation of the modified graph for finding the stationary probability of being in state 1 and state 2 is shown in Figure 7. The first derivative of the solution of this modified graph will yield the expected time that the process is in the node of interest.

Solving for the expected time in state 1 by using the topological relationship:

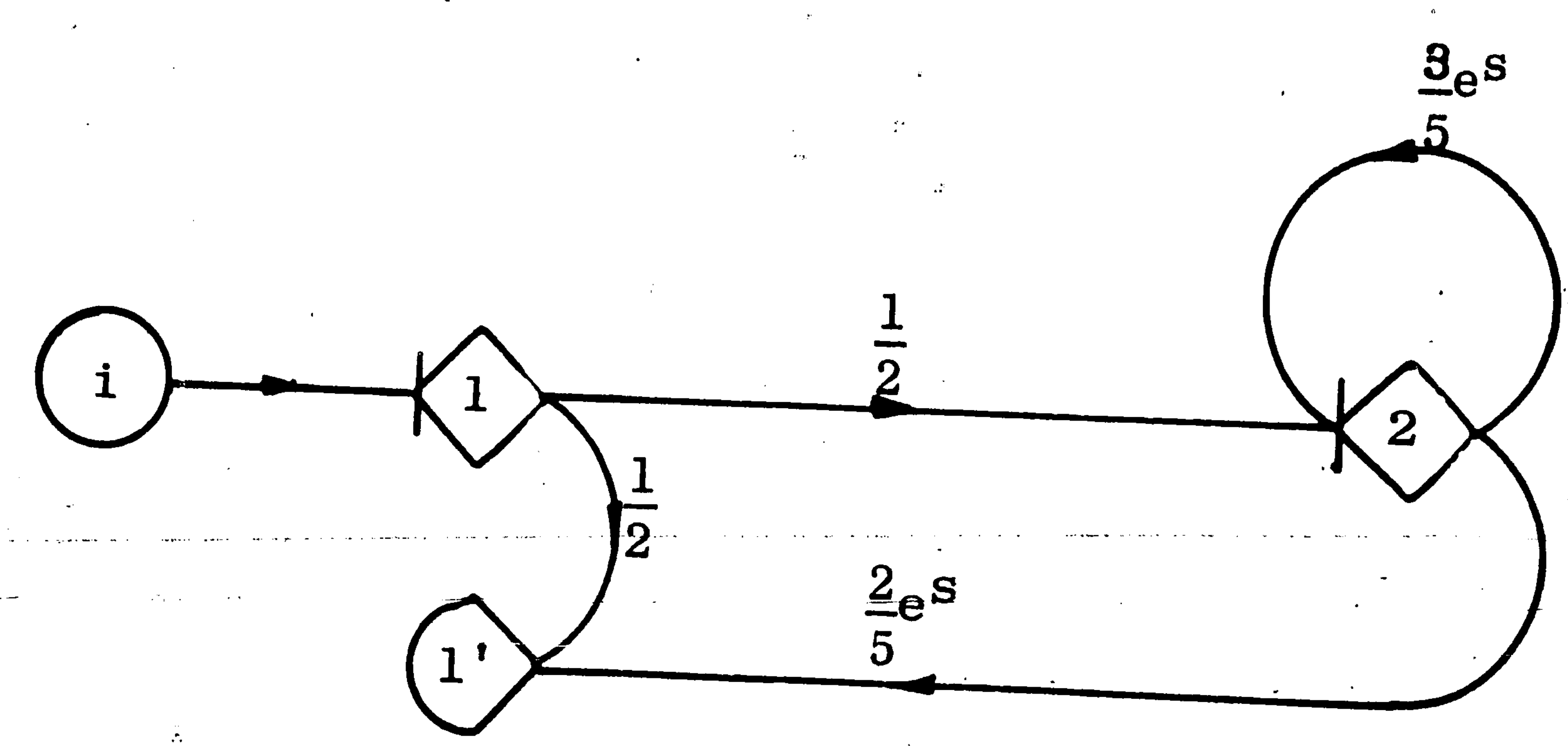
$$W_1(s) = \frac{\frac{1}{2}e^s\left[1 - \frac{3}{5}\right] + \frac{1}{5}e^s}{\left[1 - \frac{3}{5}\right]} = \frac{1}{2}e^s + \frac{1}{2}e^s$$

$$p_1 = W_1(s)\Big|_{s=0} = 1$$

$$M_1(s) = \frac{W_1(s)}{p_1} = \frac{1}{2}e^s + \frac{1}{2}e^s$$



Expected Time in State 1



Expected Time in State 2

Figure 7. GERT Representation to Find the Stationary Probability of Being in a Given State of the Toymaker Problem.

$$\mu_1 = \frac{d}{ds} M_1(s)\Big|_{s=0} = 1$$

In a similar manner, the expected time in state 2 is calculated:

$$W_2(s) = \frac{\frac{1}{2}\left[1 - \frac{3}{5}e^s\right] + \frac{1}{5}e^s}{\left[1 - \frac{3}{5}e^s\right]} = \frac{\frac{1}{2} - \frac{1}{10}e^s}{\left[1 - \frac{3}{5}e^s\right]}$$

$$p_2 = W_2(s)\Big|_{s=0} = 1$$

$$M_2(s) = \frac{W_2(s)}{p_2} = \frac{\frac{1}{2} - \frac{1}{10}e^s}{\left[1 - \frac{3}{5}e^s\right]}$$

$$\mu_2 = \left[\frac{d}{ds} M_2(s)\right]_{s=0} = \frac{5}{4}$$

The ratio of the expected time in a state of interest and the mean recurrence time will be the steady state probability of being in that state.

$$\pi_1 = \frac{\mu_1}{\mu_{rt}} = \frac{1}{\frac{9}{4}} = \frac{4}{9}$$

$$\pi_2 = \frac{\mu_2}{\mu_{rt}} = \frac{\frac{5}{4}}{\frac{9}{4}} = \frac{5}{9}$$

These values agree with those obtained by Howard (7, pp. 11) using the Z-transform method. This example shows that GERT enables one to solve for questions of interest in a Markov process.

Markov Processes With Rewards

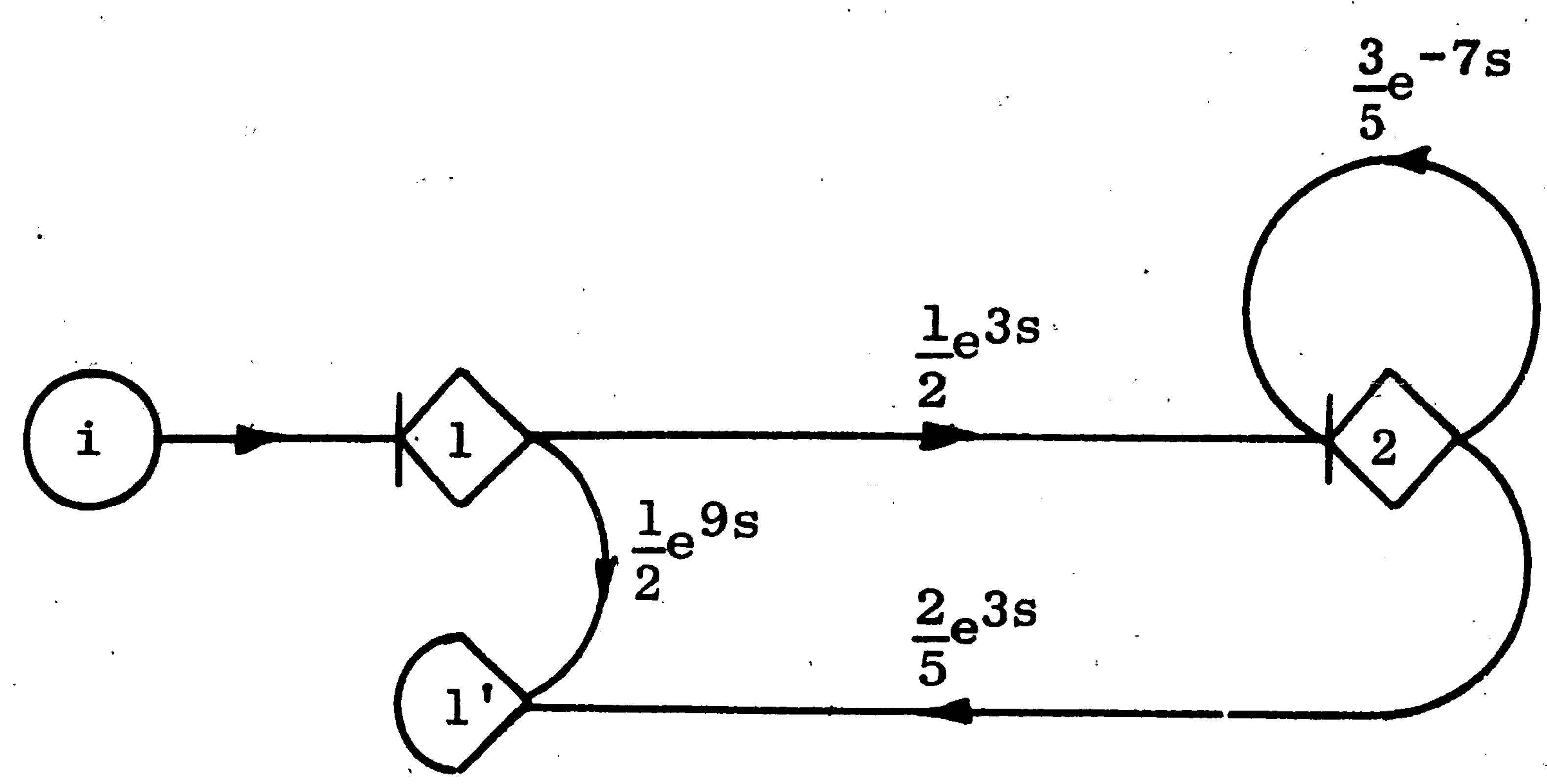
Economic Rewards may be associated with state transitions. The rewards could be in terms of dollars, voltage levels, units of production or any other physical quantity relevant to the problem. The rewards could equally well be constructed as penalties.

The Markov process generates a sequence of rewards as it makes transitions from state to state. The reward is, thus, a random variable with a probability distribution governed by the probabilistic relations of the Markov process. To investigate the problem of total reward or expected earnings in greater detail, a reward structure can be added to the toymaker's problem. The reward matrix denotes the expected reward from state i to state j and is given as follows:

$$R = \left[r_{ij} \right] = \begin{array}{c} \\ R_1 \\ R_2 \end{array} \begin{array}{c} \begin{array}{cc} R_1 & R_2 \end{array} \\ \begin{bmatrix} 9 & 3 \\ 3 & -7 \end{bmatrix} \end{array}$$

Since each transition reward is directly related to the transition time, the MGF of the time function can be replaced by the MGF of the reward function in the GERT representation to calculate the mean return of the system with rewards.

With the aid of the GERT Representation of the reward structure shown in Figure 8, the mean return of the system can be calculated the same way as the mean recurrence time was done before.



GERT Representation to Calculate Mean Return of
The Toymaker Problem
Figure 8

Therefore, using the topological equation to solve for the transmittance between state 1 and state 1':

$$W_e(s) = \frac{\frac{1}{2}e^{9s}\left[1 - \frac{3}{5}e^{-7s}\right] + \frac{1}{5}e^{6s}}{\left[1 - \frac{3}{5}e^{-7s}\right]} = \frac{\frac{1}{2}e^{9s} - \frac{3}{10}e^{2s} + \frac{1}{5}e^{6s}}{\left[1 - \frac{3}{5}e^{-7s}\right]}$$

$$p_e = W_e(s)\Big|_{s=0} = 1$$

$$M_e(s) = \frac{W_e(s)}{p_e} = \frac{\frac{1}{2}e^{9s} - \frac{3}{10}e^{2s} + \frac{1}{5}e^{6s}}{\left[1 - \frac{3}{5}e^{-7s}\right]}$$

$$\mu_r = \left[\frac{d}{ds} M_e(s)\right]_{s=0} = \frac{9}{4}$$

The gain of the system can now be defined as the ratio of the mean return to the mean recurrence time. This quantity is interpreted to be the rate of return of the toymaker's process.

$$G = \frac{\mu_r}{\mu_{rt}} = \frac{\frac{9}{4}}{\frac{9}{4}} = 1$$

Introduction of Alternatives

The toymaker has other courses of action open to him that change the probabilities and rewards governing the process. For example, the toymaker could spend money on marketing his product which would tend to increase his probability of staying in state 1, but at the same time reduce profits. He could also spend money on research and development on a new product which would tend to decrease his probability of remaining in state 2, but also reduce profits. The alternatives, as they apply to the toymaker, can therefore be summarized in vector form as follows:

$$p_{1j}^{1} = [0.5 \quad 0.5]; \qquad r_{1j}^{1} = [9 \quad 3]$$
$$p_{1j}^{2} = [0.8 \quad 0.2]; \qquad r_{1j}^{2} = [4 \quad 4]$$

$$p_{2j}^{1} = [0.4 \quad 0.6]; \qquad r_{2j}^{1} = [3 \quad -7]$$
$$p_{2j}^{2} = [0.7 \quad 0.3]; \qquad r_{2j}^{2} = [1 \quad -19]$$

The notation p_{ij}^{k} and r_{ij}^{k} represent the probabilities and rewards from state i to state j given by the k^{th} alternative.

An optimal policy is that choice of alternatives which maximize expected return or gain of the entire process. It is obvious that the gain can be evaluated for each set of alternatives separately, and the set of alternatives chosen which in fact yield the largest gain. However, it does not take too many more states and sets of alternatives to render this type of analysis useless. For example, in a five state problem with three alternatives available at each state, there are $3^5 = 243$ different policies and 243 different gains to solve.

The Policy-Iteration Method

Howard (7, Chapter 4) describes the policy-iteration method which locates the optimal policy of systems that have a large number of transitions in a small number of iterations. It is composed of two parts, the value-determination operation and the policy-improvement routine.

The value-determination operation essentially is the solution of N(the number of states in the system) linear simultaneous equations to find the gain and relative values of a given policy. The relative values in question are the total expected reward that the system will earn in n moves if it starts from state i, relative to starting in state N, under a given policy.

The policy-improvement routine uses the relative values as solved above and the expected immediate reward of each state to evaluate the policy under consideration for each individual state and to suggest a more optimal policy if one is available. The new policy will have a higher gain than the old policy. Howard summarizes the policy-iteration method as follows:

> (1) The solution of the sequential decision process is reduced to solving sets of linear simultaneous equations and subsequent comparisons.

(2) Each succeeding policy found in the iteration cycle has a higher gain than the previous one.

(3) The iteration cycle will terminate on the policy that has the largest gain attainable within the realm of the problem; it will usually find this policy in a small number of iterations.

Returning once again to the toymaker problem, where there are two states and two alternatives in each state, the toymaker has four possible policies, each with associated probabilities and rewards. The toymaker would like to know which of these four policies should be followed into the indefinite future to make the average earnings per week as large as possible.

The policy iteration routine is begun by choosing as an initial policy the one that maximizes expected immediate reward in each state. The expected immediate reward can be determined by partitioning the system in such a way that only the transactions emanating from the state of interest are considered. The gain of the partition can be calculated in the same way as the gain was calculated for the system as a whole.

$$q_i^k = \frac{\mu_{r_i}^{\ k}}{\mu_{rt_i}^{\ k}}$$; where i is the state of interest and k is the alternative within that state

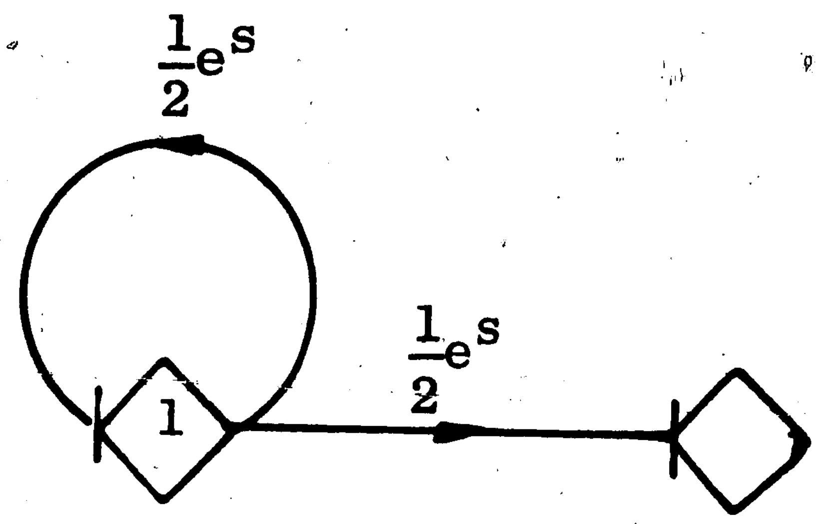
The partitioning and calculations of the expected immediate reward for the toymaker problem are detailed in Figures 9a and 9b.

Since $\left[q_1^{\ 1} = 6\right] \geq \left[q_1^{\ 2} = 4\right]$ and $\left[q_2^{\ 1} = -3\right] > \left[q_2^{\ 2} = -5\right]$, alternative 1 in states 1 and 2 is the policy that maximizes expected

STATE 1

ALTERNATIVE 1

Mean Recurrence Time



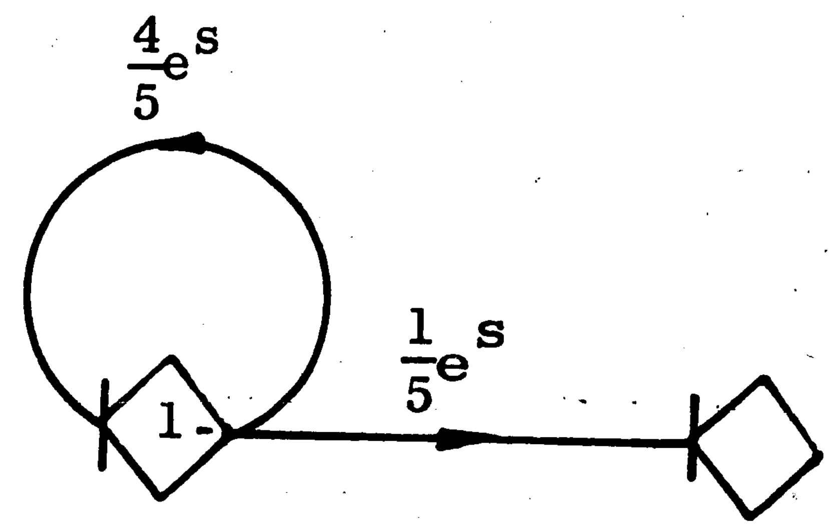
$$W(s) = \frac{\frac{1}{2}e^{s}}{\left[1 - \frac{1}{2}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{p_e} = \frac{\frac{1}{2}e^{s}}{\left[1 - \frac{1}{2}e^{s}\right]}$$

$$\mu_{rt} = \left[\frac{d}{ds} M(s)\right]_{s=0} = 2$$

ALTERNATIVE 2

Mean Recurrence Time



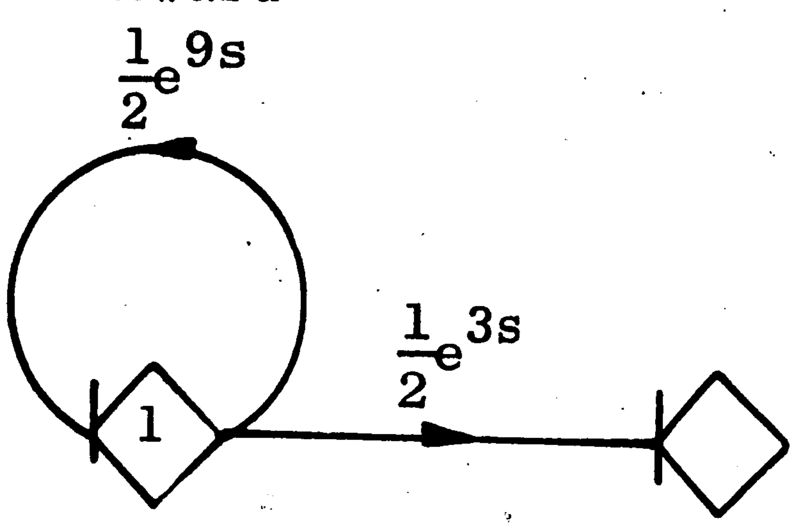
$$W(s) = \frac{\frac{1}{5}e^{s}}{\left[1 - \frac{4}{5}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{p_e} = \frac{\frac{1}{5}e^{s}}{\left[1 - \frac{4}{5}e^{s}\right]}$$

$$\mu_{rt} = \left[\frac{d}{ds} M(s)\right]_{s=0} = 5$$

ALTERNATIVE 1

Mean Reward



$$W(s) = \frac{\frac{1}{2}e^{3s}}{\left[1 - \frac{1}{2}e^{9s}\right]}$$

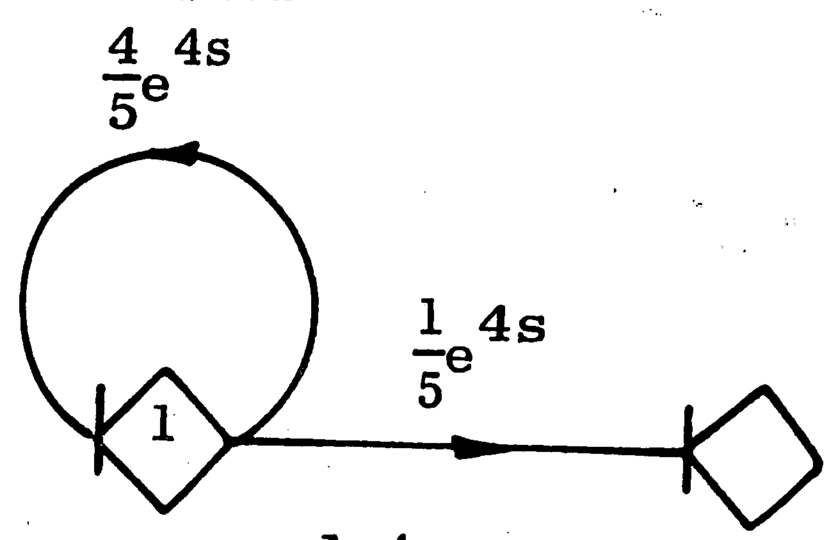
$$M(s) = \frac{W(s)}{p_e} = \frac{\frac{1}{2}e^{3s}}{\left[1 - \frac{1}{2}e^{9s}\right]}$$

$$\mu_{r} = \left[\frac{d}{ds} M(s)\right]_{s=0} = 12$$

$$\therefore q_1^{\,1} = \frac{\mu_r}{\mu_{rt}} = \frac{12}{2} = 6$$

ALTERNATIVE 2

Mean Reward



$$W(s) = \frac{\frac{1}{5}e^{4s}}{\left[1 - \frac{4}{5}e^{4s}\right]}$$

$$M(s) = \frac{W(s)}{p_e} = \frac{\frac{1}{5}e^{4s}}{\left[1 - \frac{4}{5}e^{4s}\right]}$$

$$\mu_{r} = \left[\frac{d}{ds} M(s)\right]_{s=0} = 20$$

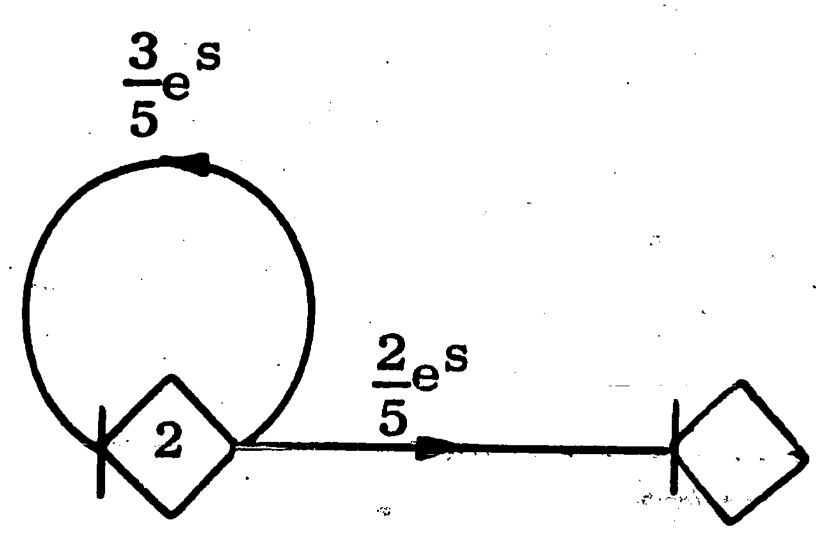
$$\therefore q_1^{\,2} = \frac{\mu_r}{\mu_{rt}} = \frac{20}{5} = 4$$

Figure 9a. GERT Representation for Calculating the Expected Immediate Reward in State 1

STATE 2

ALTERNATIVE 1

Mean Recurrence Time

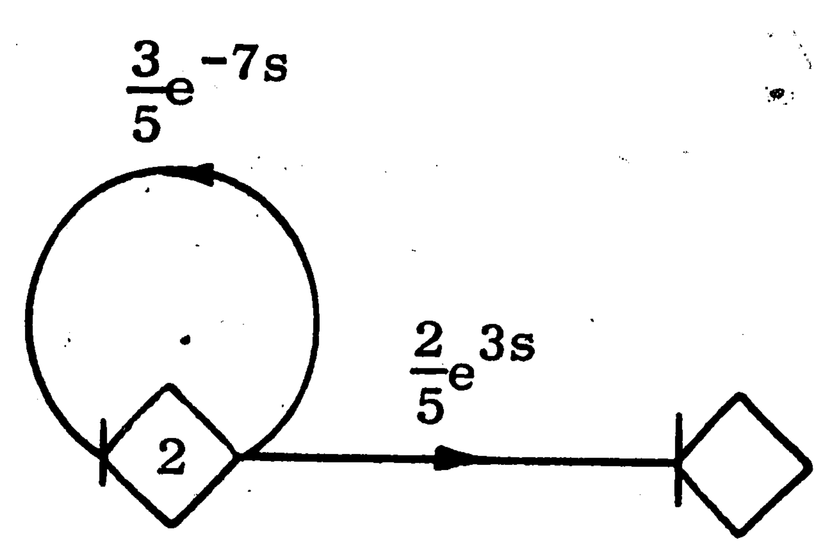


$$W(s) = \frac{\frac{2}{5}e^{s}}{\left[1 - \frac{3}{5}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{p} = W(s)$$

$$\mu_{rt} = \left[\frac{d}{ds} M(s)\right]_{s=0} = \frac{5}{2}$$

Mean Reward



$$W(s) = \frac{\frac{2}{5}e^{3s}}{\left[1 - \frac{3}{5}e^{-7s}\right]}$$

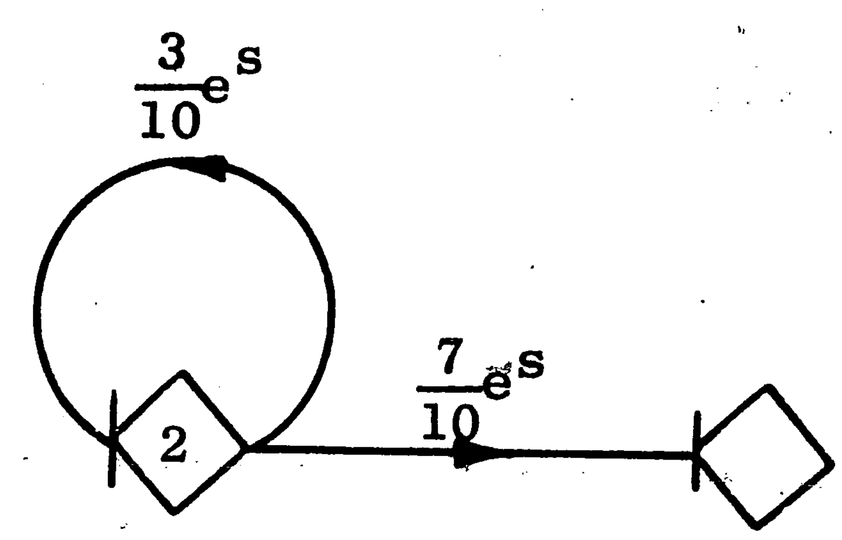
$$M(s) = \frac{W(s)}{p} = W(s)$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = -\frac{15}{2}$$

$$\therefore q_2^{\,1} = \frac{\mu_r}{\mu_{rt}} = \frac{-\frac{15}{2}}{\frac{5}{2}} = -3$$

ALTERNATIVE 2

Mean Recurrence Time

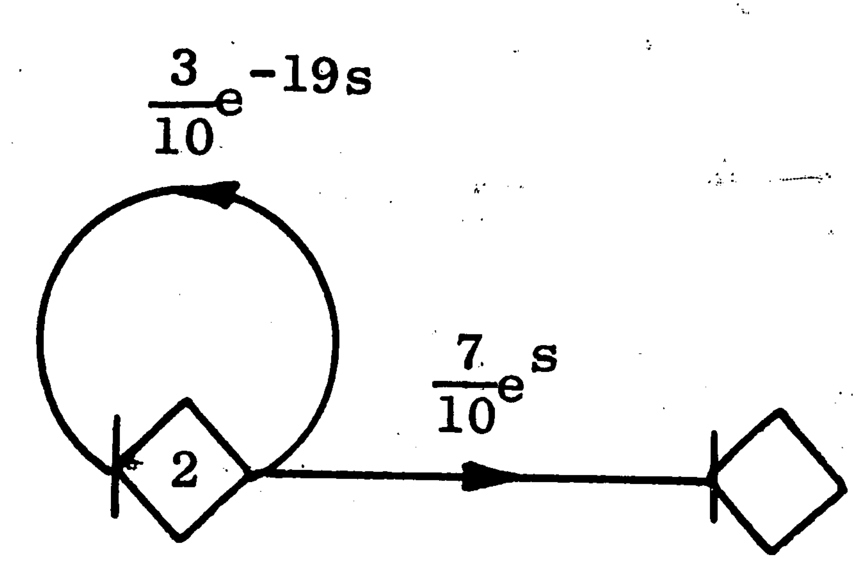


$$W(s) = \frac{\frac{7}{10}e^{s}}{\left[1 - \frac{3}{10}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{p} = W(s)$$

$$\mu_{rt} = \left[\frac{d}{ds} M(s)\right]_{s=0} = \frac{10}{7}$$

Mean Reward



$$W(s) = \frac{\frac{7}{10}e^{s}}{\left[1 - \frac{3}{10}e^{-19s}\right]}$$

$$M(s) = \frac{W(s)}{p} = W(s)$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = -\frac{50}{7}$$

$$\therefore q_2^{\,2} = \frac{\mu_r}{\mu_{rt}} = \frac{-\frac{50}{7}}{\frac{10}{7}} = -5$$

Figure 9b. GERT Representation for Calculating the Expected Immediate Reward in State 2.

immediate reward. The system gain of this choice of alternatives, $k = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, has been previously calculated on page 24 and determined to be 1.

The first phase of Howard's policy iteration routine is the value-determination operation. This operation evaluates the initial policy chosen above by solving the N simultaneous equations:

$$g + v_i = q_i + \sum_{j=1}^{N} p_{ij} v_j \; ; \quad i = 1, 2, \ldots N$$

where g is the gain of the system with the chosen policy.
q is the expected immediate reward of state i with the chosen policy.
p is the transition probability of going to state j from state i with the chosen policy.
and v is the relative value of starting in state i rather than state N.

Since v_i is relative to v_N for all i, v_N can be set equal to zero. There are therefore N simultaneous equations and N $(g, v_1, \ldots v_{N-1})$ unknowns.

$$g + v_1 = 6 + \frac{1}{2}v_1 + \frac{1}{2}v_2$$

$$g + v_2 = -3 + \frac{2}{5}v_1 + \frac{3}{5}v_2$$

Setting $v_2 = 0$ and solving these equations, the following values are obtained:

$g = 1$ $\qquad v_1 = 10$ $\qquad v_2 = 0$

The second phase of Howard's policy iteration routine is the policy-improvement routine. For each state i, the alternative k that maximizes the test quantity:

$$q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j$$

using the relative values determined under the old policy is an improvement.

State	Alternative	Test Quantity
1	1	$6 + \frac{1}{2}(10) + \frac{1}{2}(0) = 11$
	2	$4 + \frac{4}{5}(10) + \frac{1}{5}(0) = 12$
2	1	$-3 + \frac{2}{5}(10) + \frac{3}{5}(0) = 1$
	2	$-5 + \frac{7}{10}(10) + \frac{3}{10}(0) = 2$

The policy-improvement routine reveals that the second alternative in each state produces a higher value of the test quantity than does the first alternative. Thus the policy composed of the second alternative in each state will have a higher gain than the original policy. The procedure must be continued to find whether a more optimal policy is available.

Re-iterating back to the value-determination to evaluate the new policy:

$$g + v_1 = 4 + \frac{4}{5}v_1 + \frac{1}{5}v_2$$
$$g + v_2 = -5 + \frac{7}{10}v_1 + \frac{3}{10}v_2$$

with $v_2 = 0$, the results of the value determination operation are

$$g = 2 \qquad v_1 = 10 \qquad v_2 = 0$$

The policy-improvement routine must be entered again, but since the relative values are coincidentally the same as those for the previous iteration, the test quantity calculations are merely repeated. The policy, $k = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is found once more, and, since the same policy has been found twice in succession, the optimal policy with a corresponding rate of return of 2 has been found.

A minimization problem can be evaluated in a similar manner by noting that the policy can be improved by the alternatives for each state which minimize the test-quantity.

CHAPTER IV

THE TAXICAB OPERATION

A further example of sequential decision process is presented in this chapter to show how GERT may be applied to a variety of problems. The taxicab operation is a more complex discrete-time Markov process, but the optimization technique is the same as the simpler toymaker problem.

> Consider the problem of the taxicab driver whose territory encompasses three towns, A, B, and C. If he is in town A, he has three alternatives:
>
> (1) He can cruise in the hope of picking up a passenger by being hailed.
>
> (2) He can drive to the nearest cab stand and wait in line.
>
> (3) He can pull over and wait for a radio call.
>
> If he is in town C, he has the same three alternatives, but if he is in town B, the last alternative is not present because there is no radio cab service in that town. For a given town and given alternative, there is a probability that the next trip will go to each of the towns, A, B, and C, and a corresponding reward in monetary units associated with each such trip. This reward represents the income from the trip after all necessary expenses have been deducted.

The probabilities of transition and the rewards of this problem depend upon the alternative because different customer population will be encountered under each alternative. If the towns, A, B, and C are identified as states 1, 2, and 3, respectively, then Figure 10 shows the GERT network of the operation. In addition, the probability and reward structure is given as follows:

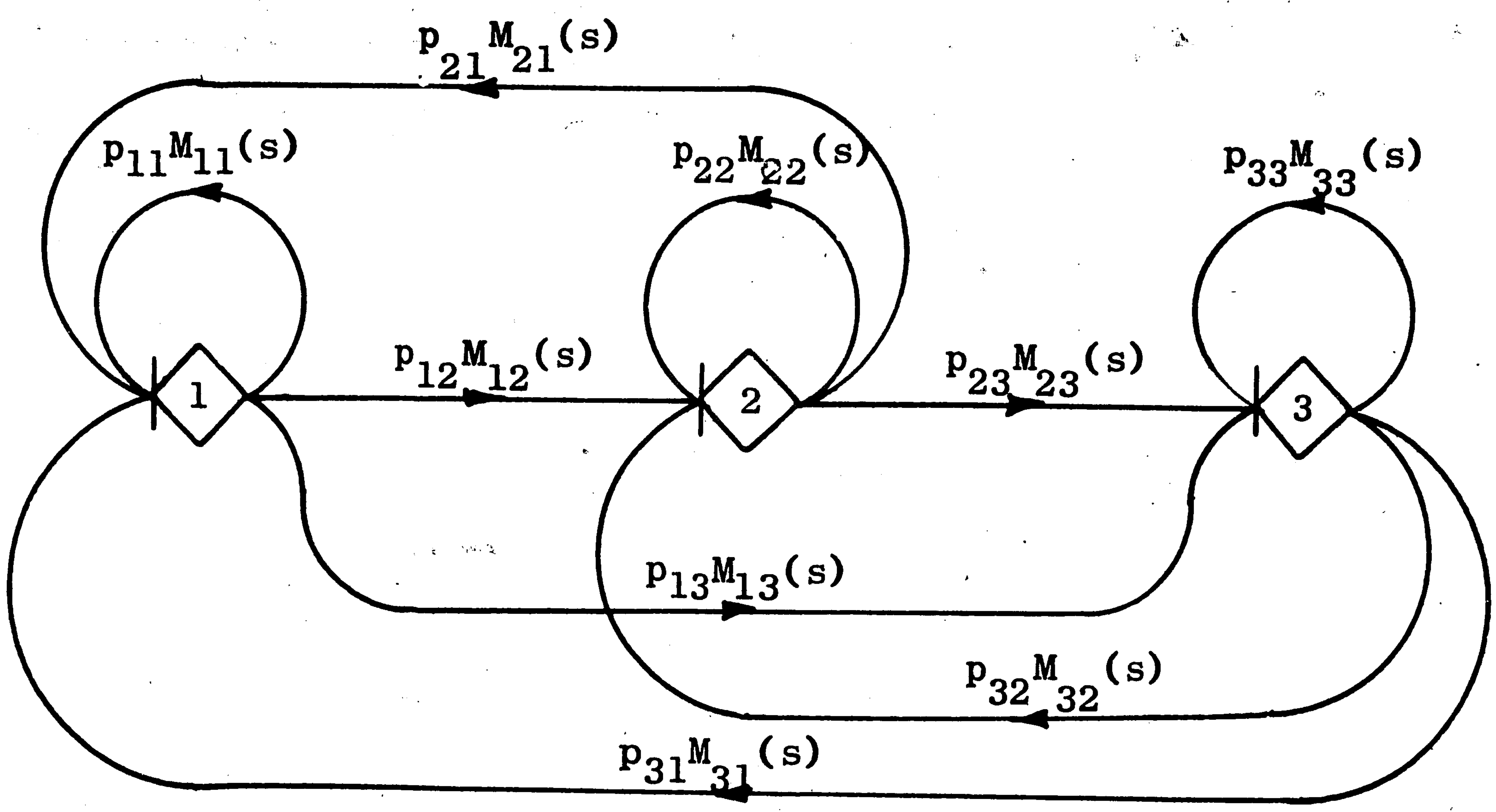


Figure 10. GERT Representation of the Taxicab Operation

$$p_{1j}^{1} = [0.5 \quad 0.25 \quad 0.25] \qquad r_{1j}^{1} = [10 \quad 4 \quad 8]$$
$$p_{1j}^{2} = [0.0625 \quad 0.75 \quad 0.1875] \qquad r_{1j}^{2} = [8 \quad 2 \quad 4]$$
$$p_{1j}^{3} = [0.25 \quad 0.125 \quad 0.625] \qquad r_{1j}^{3} = [4 \quad 6 \quad 4]$$

$$p_{2j}^{1} = [0.5 \quad 0 \quad 0.5] \qquad r_{2j}^{1} = [14 \quad 0 \quad 18]$$
$$p_{2j}^{2} = [0.0625 \quad 0.875 \quad 0.0625] \qquad r_{2j}^{2} = [8 \quad 16 \quad 8]$$

$$p_{3j}^{1} = [0.25 \quad 0.25 \quad 0.5] \qquad r_{3j}^{1} = [10 \quad 2 \quad 8]$$
$$p_{3j}^{2} = [0.125 \quad 0.75 \quad 0.125] \qquad r_{3j}^{2} = [6 \quad 4 \quad 2]$$
$$p_{3j}^{3} = [0.75 \quad 0.0625 \quad 0.1875] \qquad r_{3j}^{3} = [4 \quad 0 \quad 8]$$

The gain for each alternative at each individual state is calculated using the relationship:

$$q_i^k = \frac{\mu_{r_i}{}^k}{\mu_{rt_i}{}^k}$$

The calculations for states 1, 2, and 3 are summarized in Figure 11a, 11b, and 11c. To begin the iteration process, the policy that maximizes the expected immediate reward, or has the largest gain for each individual state, is chosen.

The policy that maximizes immediate expected reward is found to be $k = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Howard's policy-iteration routine described in the previous chapter can now be entered to evaluate this initial policy and locate the optimal policy.

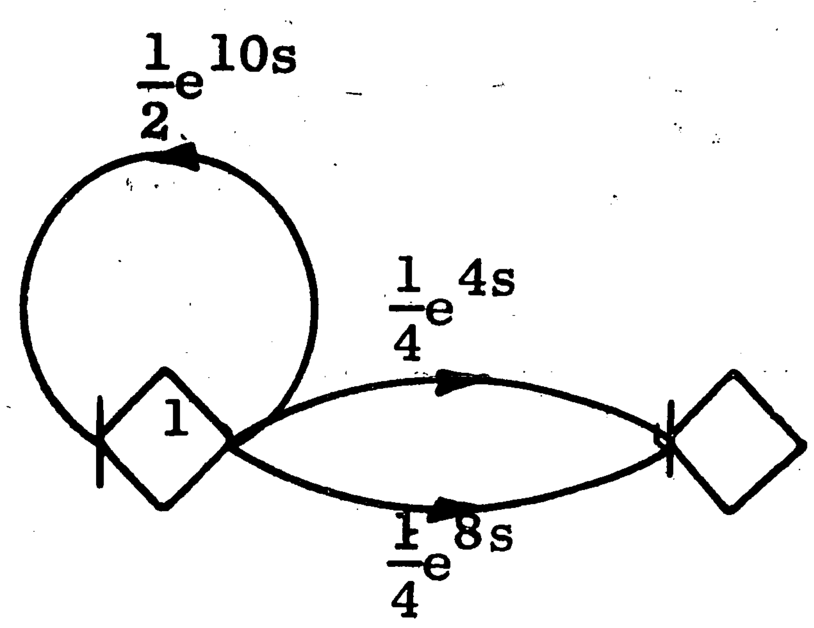
ALTERNATIVE 1

$$M(s) = \frac{W(s)}{p} = \frac{\frac{1}{4}e^{4s} + \frac{1}{4}e^{8s}}{\left[1 - \frac{1}{2}e^{10s}\right]}$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = 16$$

similarly, $\mu_{rt} = 2$

$$\therefore q_1^{\,1} = \frac{\mu_r}{\mu_{rt}} = \frac{16}{2} = 8$$



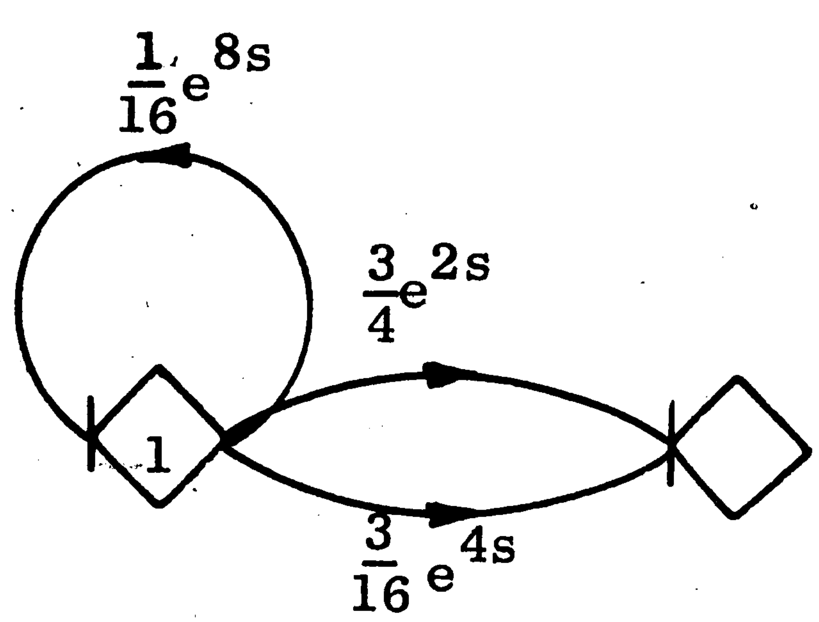
ALTERNATIVE 2

$$M(s) = \frac{W(s)}{p} = \frac{\frac{3}{4}e^{2s} + \frac{3}{16}e^{4s}}{\left[1 - \frac{1}{16}e^{10s}\right]}$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = \frac{46}{15}$$

similarly, $\mu_{rt} = \frac{16}{15}$

$$\therefore q_1^{\,2} = \frac{\mu_r}{\mu_{rt}} = \frac{46}{15}\Big/\frac{16}{15} = 2.75$$



ALTERNATIVE 3

$$M(s) = \frac{W(s)}{p} = \frac{\frac{1}{8}e^{6s} + \frac{5}{8}e^{4s}}{\left[1 - \frac{1}{4}e^{4s}\right]}$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = \frac{17}{3}$$

similarly, $\mu_{rt} = \frac{4}{3}$

$$\therefore q_1^{\,3} = \frac{\mu_r}{\mu_{rt}} = \frac{17}{3}\Big/\frac{4}{3} = 4.25$$

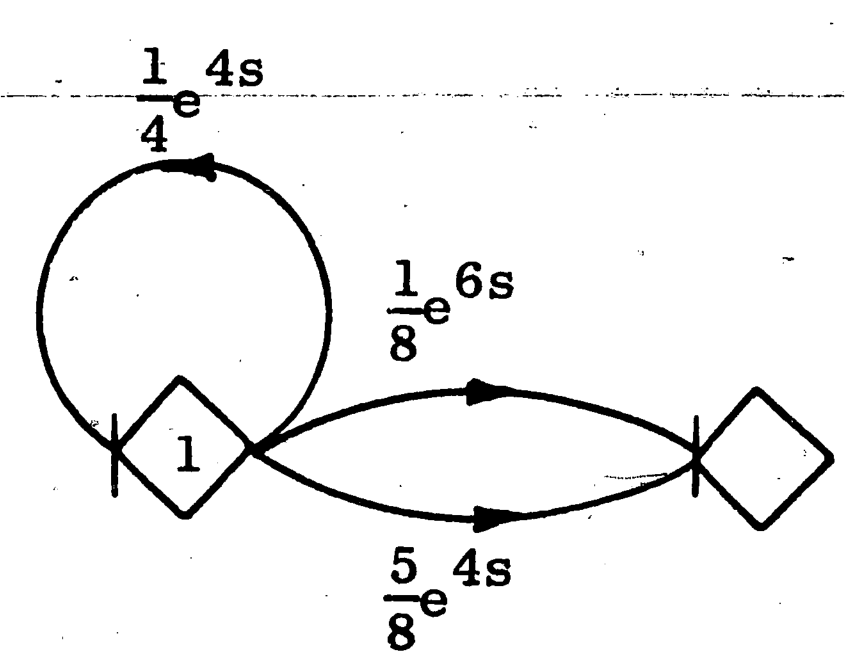
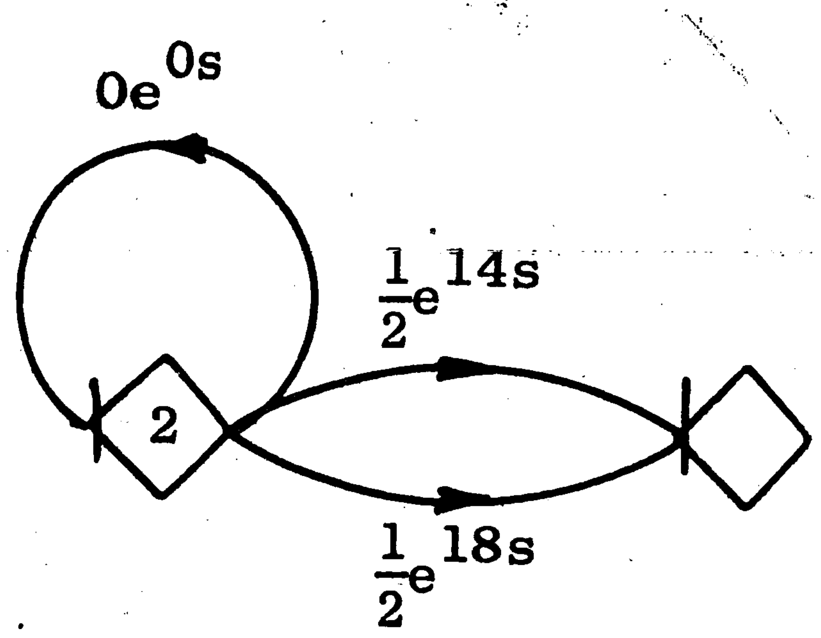


Figure 11a. GERT Representation to Calculate Expected Immediate Reward in State 1.

ALTERNATIVE 1



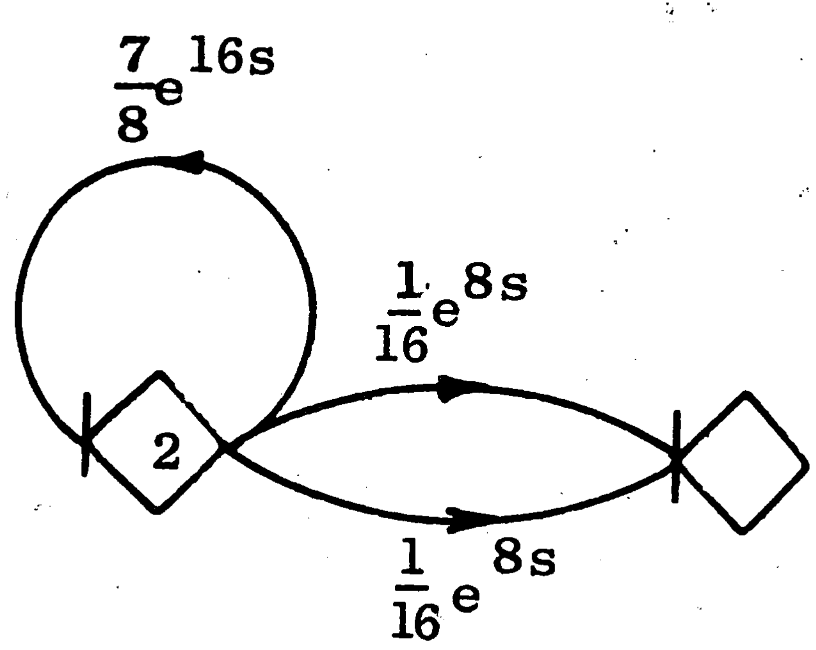
$$M(s) = \frac{W(s)}{p} = \frac{1}{2}e^{14s} + \frac{1}{2}e^{18s}$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = 16$$

similarly, $\mu_{rt} = 1$

$$\therefore q_2^1 = \frac{\mu_r}{\mu_{rt}} = \frac{16}{1} = 16$$

ALTERNATIVE 2



$$M(s) = \frac{W(s)}{p} = \frac{\frac{1}{16}e^{8s} + \frac{1}{16}e^{8s}}{\left[1 - \frac{7}{8}e^{16s}\right]}$$

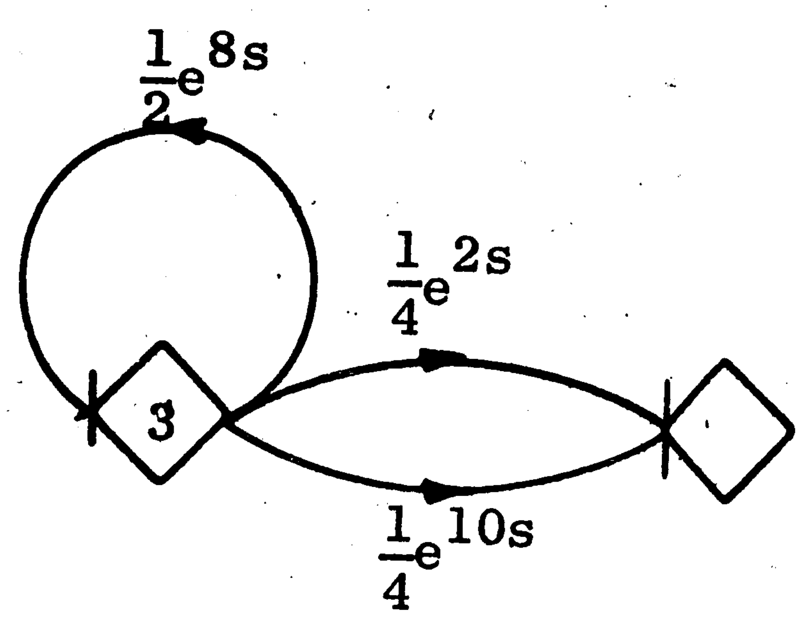
$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = 120$$

similarly, $\mu_{rt} = 8$

$$\therefore q_2^2 = \frac{\mu_r}{\mu_{rt}} = \frac{120}{8} = 15$$

Figure 11b. GERT Representation to Calculate Expected Immediate Reward in State 2

ALTERNATIVE 1



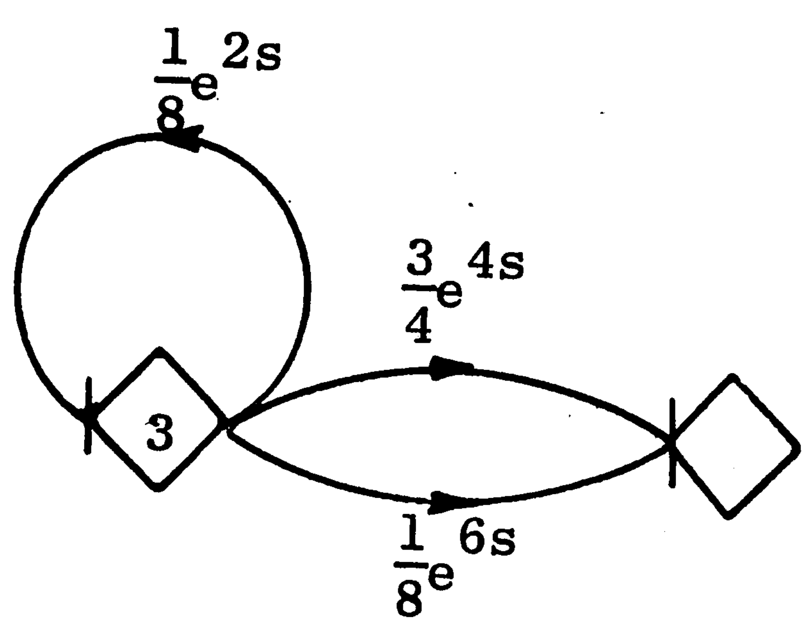
$$M(s) = \frac{W(s)}{p} = \frac{\frac{1}{4}e^{2s} + \frac{1}{4}e^{10s}}{\left[1 - \frac{1}{2}e^{8s}\right]}$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = 14$$

similarly, $\mu_{rt} = 2$

$$\therefore q_3^{\,1} = \frac{\mu_r}{\mu_{rt}} = \frac{14}{2} = 7$$

ALTERNATIVE 2



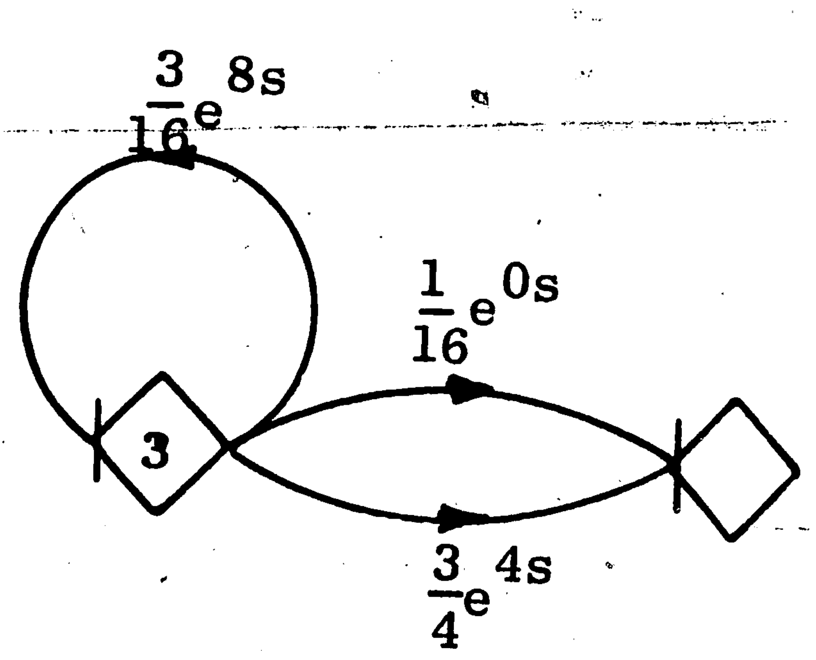
$$M(s) = \frac{W(s)}{p} = \frac{\frac{3}{4}e^{4s} + \frac{1}{8}e^{6s}}{\left[1 - \frac{1}{8}e^{2s}\right]}$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = \frac{32}{7}$$

similarly, $\mu_{rt} = \frac{8}{7}$

$$\therefore q_3^{\,2} = \frac{\mu_r}{\mu_{rt}} = \frac{32}{7}\Big/\frac{8}{7} = 4$$

ALTERNATIVE 3



$$M(s) = \frac{W(s)}{p} = \frac{\frac{1}{16} + \frac{3}{4}e^{4s}}{\left[1 - \frac{3}{16}e^{8s}\right]}$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = \frac{72}{13}$$

similarly, $\mu_{rt} = \frac{15}{13}$

$$\therefore q_3^{\,3} = \frac{\mu_r}{\mu_{rt}} = \frac{72}{13}\Big/\frac{15}{13} = 4.8$$

Figure 11c. GERT Representation to Calculate Expected Immediate Reward in State 3.

CHAPTER 5

APPLICATIONS IN A MORE GENERAL CLASS OF PROBLEMS

The techniques of GERT can be employed whenever the moment generating function of the time distribution from node to node is defined. The MGF is well defined for many familiar and representative discrete and continuous distributions such as the normal, exponential and Poisson Distributions, to name a few, as well as the discrete constant time distribution that has been used in the previous examples. GERT, therefore, is a flexible and powerful tool to be used in analyzing a more general class of problems. In particular, the policy-iteration technique can be used to solve the continuous time Markov decision problem in the same straightforward manner as the discrete-time problem was solved.

This application can be demonstrated by returning once again to the toymaker problem, and modifying the transition times from state to state, such that they are taken from a continuous distribution. The probability matrix used in the original problem is retained. To pose a hypothetical example let the distribution of the time to traverse the path from state i to state j, $f_{ij}(t)$ and its associated MGF be defined for every path as follows:

$$f_{11}(x) = 1 \qquad x = 1$$

$$M_{11}(s) = e^{s} \qquad \text{[constant]}$$

$$f_{12}(x) = ae^{-ax} \qquad x \geq 0$$

$$M_{12}(s) = (1 - \frac{s}{a})^{-1} \qquad \text{[exponential, a=1 week]}$$

$$f_{21}(x) = ae^{-ax} \qquad x \geq 0$$

$$M_{21}(s) = (1 - \frac{s}{a})^{-1} \qquad \left[\text{exponential, a=1 week}\right]$$

$$f_{22}(x) = 1 \qquad x = 1$$

$$M_{22}(s) = e^{s} \qquad \left[\text{constant}\right]$$

The problem is described graphically in Figure 12 using the new MGF to define the transmittances from state to state.

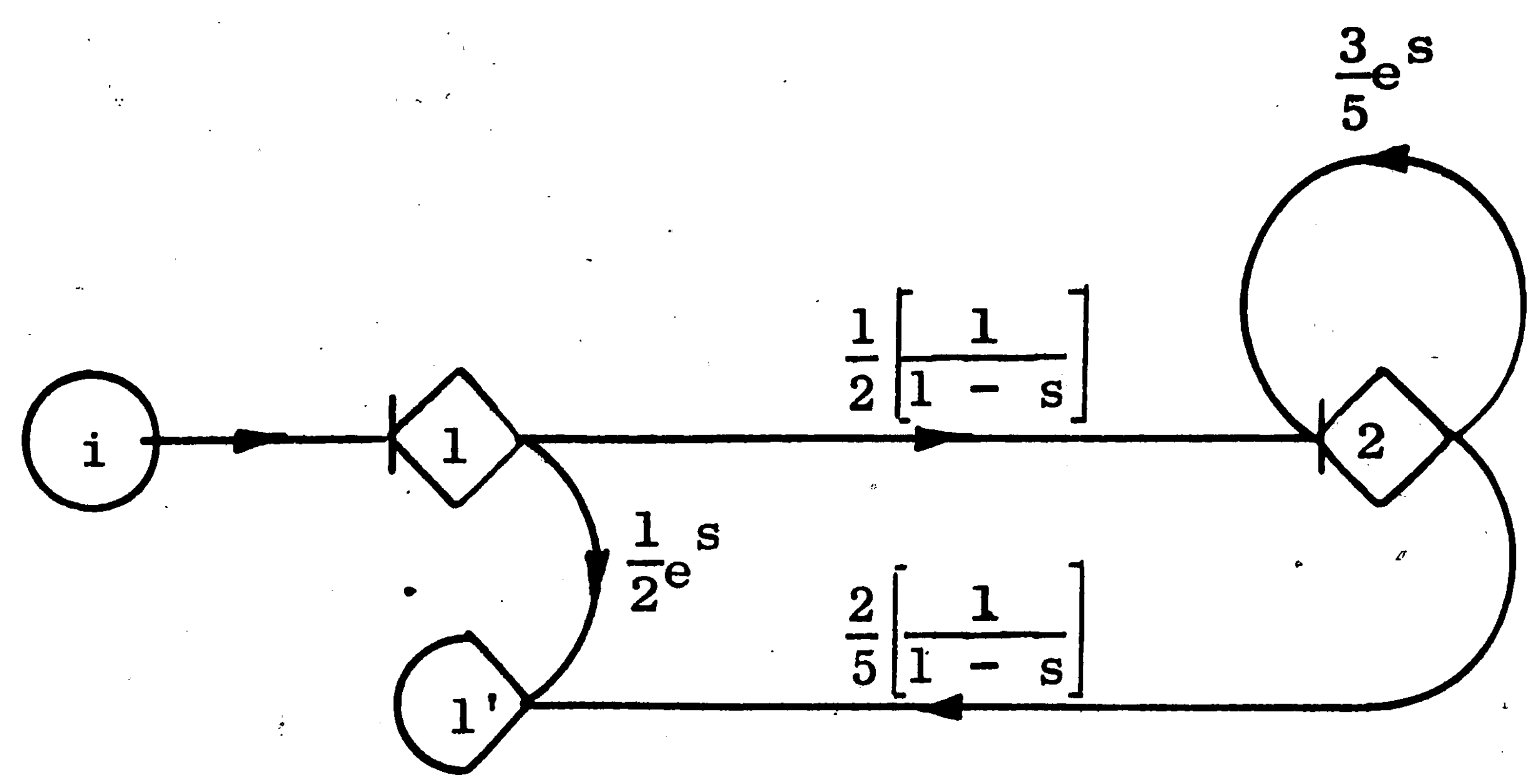


Figure 12. GERT Representation to Calculate Mean Recurrence Time

The equivalent transmittance from state 1 to state 1' of the modified toymaker problem shown graphically in Figure 12 is solved in precisely the same way as the original problem by the topological equation:

$$W_e(s) = \frac{\frac{1}{2}e^{s}\left[1 - \frac{3}{5}e^{s}\right] + \frac{1}{5}\left[\frac{1}{(1 - s)^2}\right]}{\left[1 - \frac{3}{5}e^{s}\right]}$$

$$p_e = W_e(s)\Big|_{s=0} = 1$$

$$M_e(s) = \frac{W_e(s)}{p_e} = W_e(s)$$

The mean recurrence time is therefore:

$$\mu_{rt} = \left[\frac{d}{ds}M_e(s)\right]_{s=0} = \frac{9}{4}$$

The original reward structure, $R = \begin{bmatrix} 9 & 3 \\ 3 & -7 \end{bmatrix}$ can also be incorporated into the modified toymaker problem, transforming the time transition into reward transition.

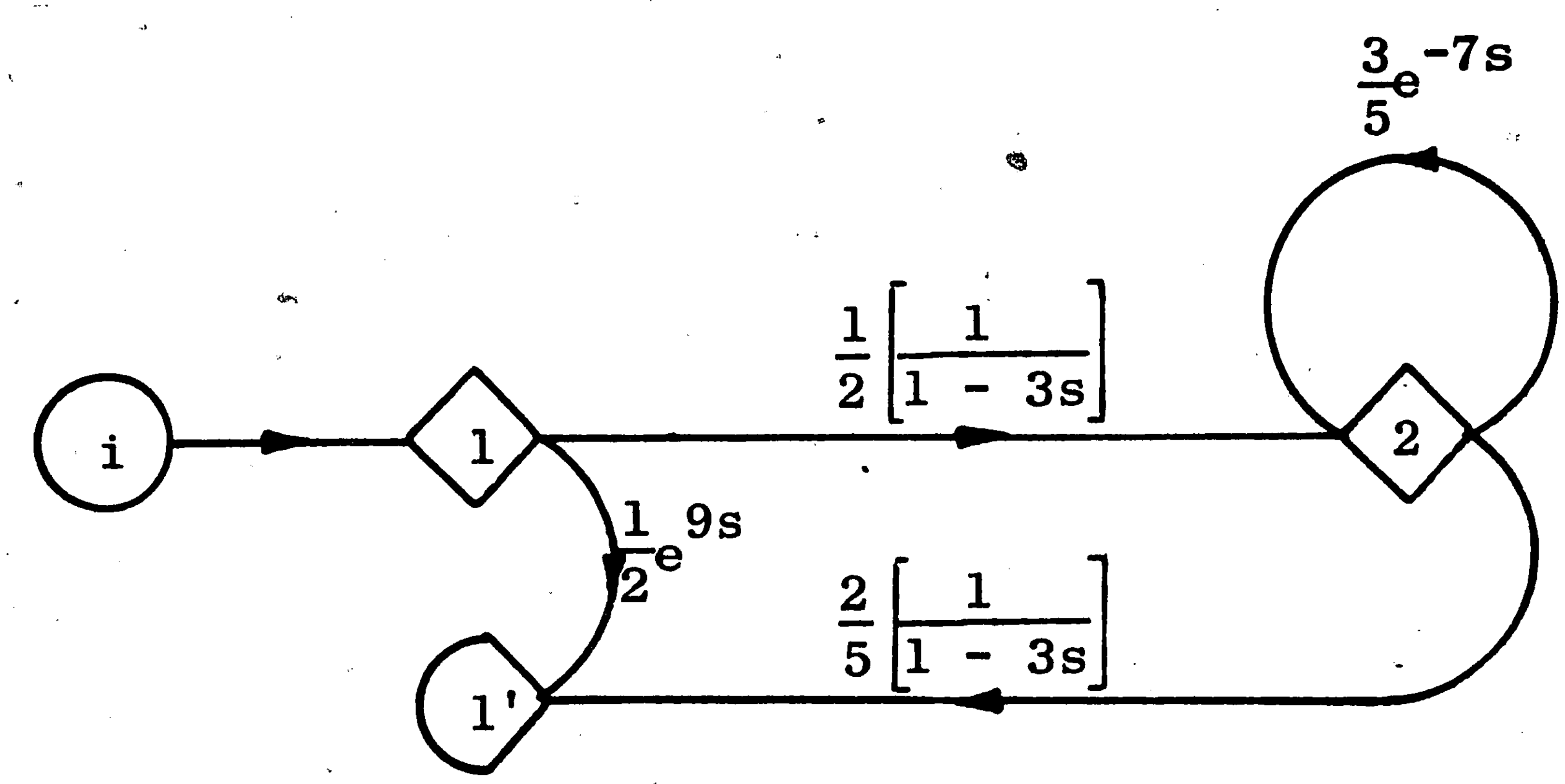


Figure 13. GERT Representation to Calculate Mean Return

Using Figure 13, the mean return is calculated as follows:

$$W_e(s) = \frac{\frac{1}{2}e^{9s}\left[1 - \frac{3}{5}e^{-7s}\right] + \frac{1}{5}\left[\frac{1}{(1-3s)^2}\right]}{\left[1 - \frac{3}{5}e^{-7s}\right]}$$

$$p_e = W_e(s)\Big|_{s=0} = 1$$

$$M_e(s) = \frac{W_e(s)}{p_e} = W_e(s)$$

$$\mu_r = \left[\frac{d}{ds}M_e(s)\right]_{s=0} = \frac{9}{4}$$

The rate of return or gain of the above system can now be calculated

$$G = \frac{\mu_r}{\mu_{rt}} = \frac{9}{4} / \frac{9}{4} = 1$$

The set of alternatives that was introduced into the original toymaker's problem on page 24 can also be applied to this new problem. These alternatives and the calculation of expected immediate rate of return (gain provided by an individual state) are denoted in Figure 14a and 14b.

Using the calculated gains of the expected immediate reward for each individual state and each alternative, the policy iteration technique can assume the role of finding the optimal policy. The initial policy is that policy which maximizes immediate reward and is found to be $k = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with corresponding $q = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$.

Sensitivity Analysis of the Policy-Iteration Technique

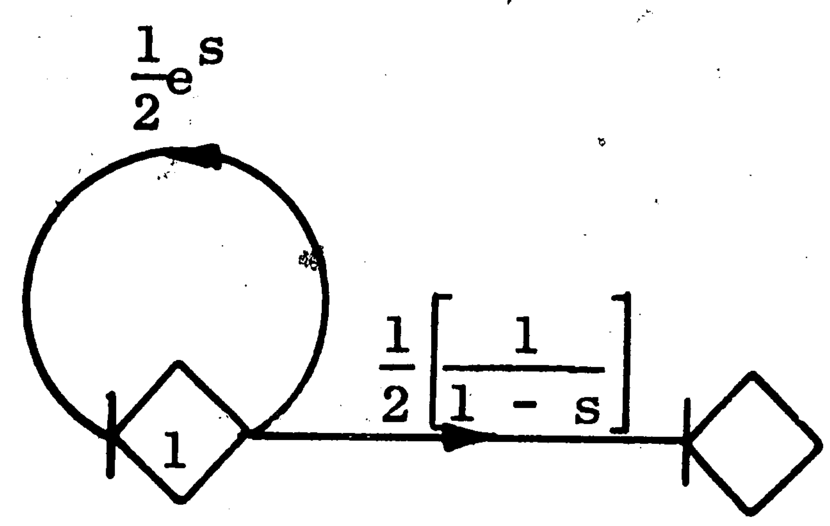
A sensitivity analysis of the policy-iteration technique is provided by exploring the changes in the system gain of the modified toymaker problem as the reward structure is varied. More specifically, Figure 15 is a graph of the rate of return of each of the four possible policies as r_{11}^{1} is varied from 9 units to 17 units.

Obviously, the rate of return for policies $k = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $k = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ remains constant since r_{11}^{1} is not included in the calculations for determining the gain for these policies. However, the rate of return for policies $k = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $k = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ increases as r_{11}^{1} increases. Further, policy $k = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ becomes the optimal policy when

STATE 1

ALTERNATIVE 1

Mean Recurrence Time

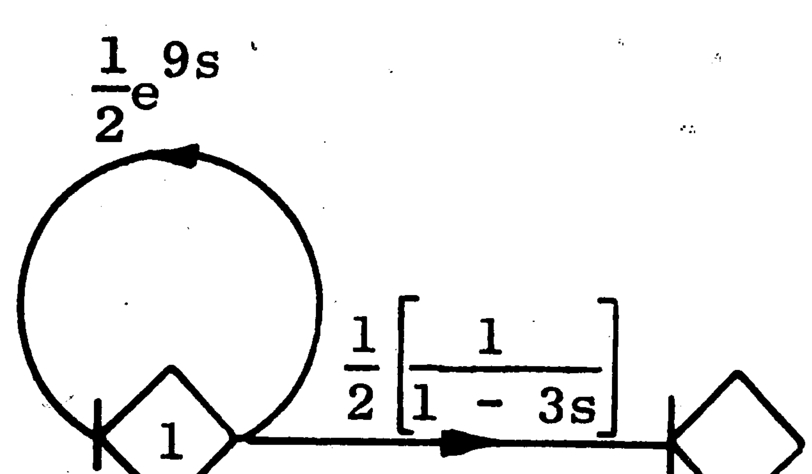


$$W(s) = \frac{\frac{1}{2}\left[\frac{1}{1 - s}\right]}{\left[1 - \frac{1}{2}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_{rt} = \left[\frac{d}{ds}M(s)\right]_{s=0} = 2$$

Mean Reward



$$W(s) = \frac{\frac{1}{2}\left[\frac{1}{1 - 3s}\right]}{\left[1 - \frac{1}{2}e^{9s}\right]}$$

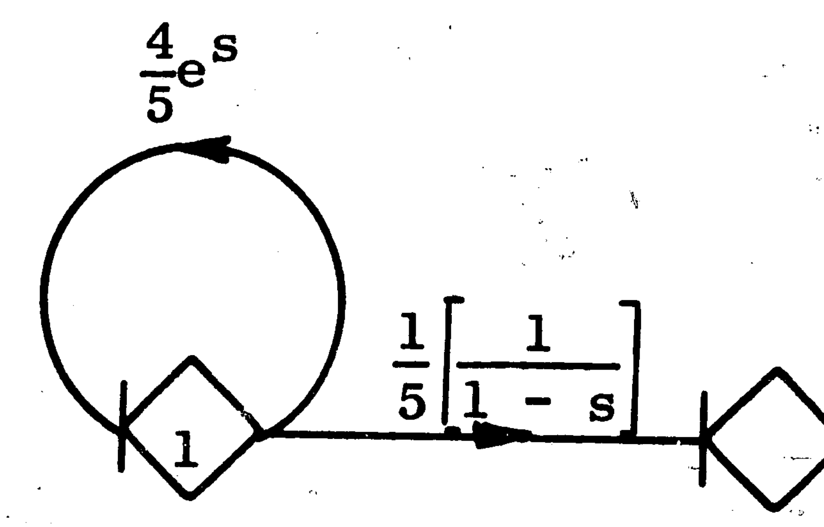
$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = 12$$

$$\therefore q_1^{\,1} = \frac{\mu_r}{\mu_{rt}} = \frac{12}{2} = 6$$

ALTERNATIVE 2

Mean Recurrence Time

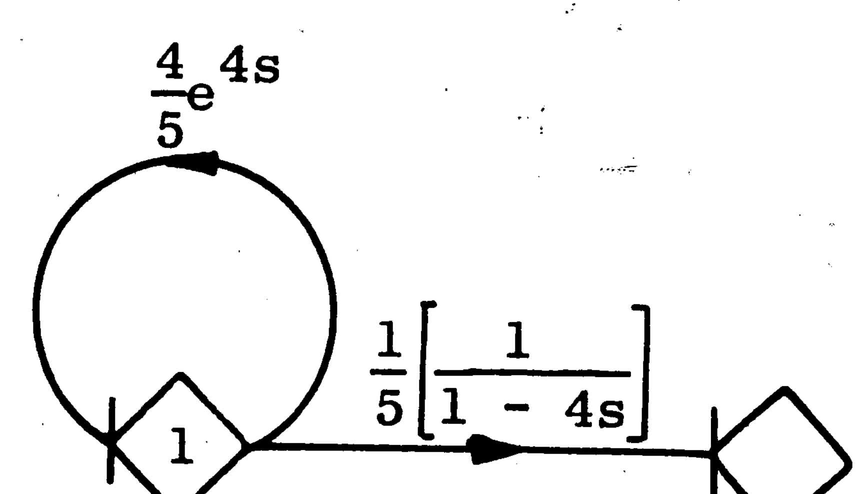


$$W(s) = \frac{\frac{1}{5}\left[\frac{1}{1 - s}\right]}{\left[1 - \frac{4}{5}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_{rt} = \left[\frac{d}{ds}M(s)\right]_{s=0} = 5$$

Mean Reward



$$W(s) = \frac{\frac{1}{5}\left[\frac{1}{1 - 4s}\right]}{\left[1 - \frac{4}{5}e^{4s}\right]}$$

$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_r = \left[\frac{d}{ds}M(s)\right]_{s=0} = 20$$

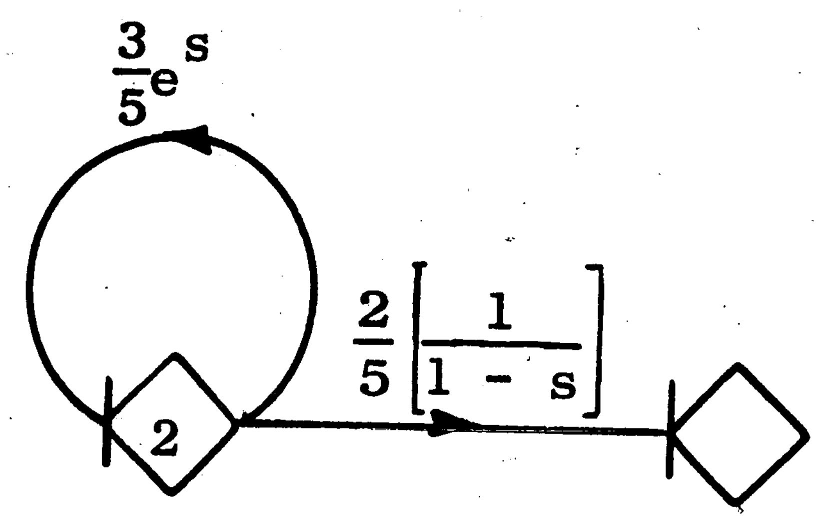
$$\therefore q_1^{\,2} = \frac{\mu_r}{\mu_{rt}} = \frac{20}{5} = 4$$

Figure 14a. GERT Representation for Calculating the Expected Immediate Reward in State 1

STATE 2

ALTERNATIVE 1

Mean Recurrence Time

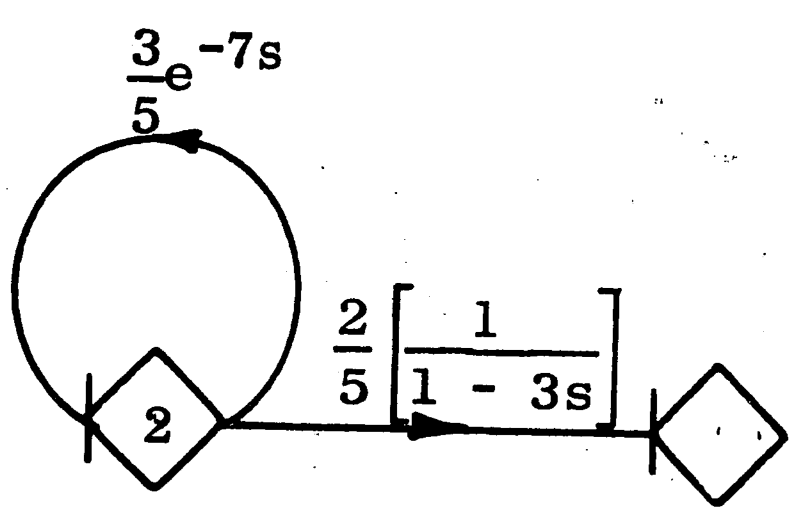


$$W(s) = \frac{\frac{2}{5}\left[\frac{1}{1-s}\right]}{\left[1 - \frac{3}{5}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_{rt} = \left[\frac{d}{ds}M(s)\right]_{s=0} = \frac{5}{2}$$

Mean Reward



$$W(s) = \frac{\frac{2}{5}\left[\frac{1}{1-3s}\right]}{\left[1 - \frac{3}{5}e^{-7s}\right]}$$

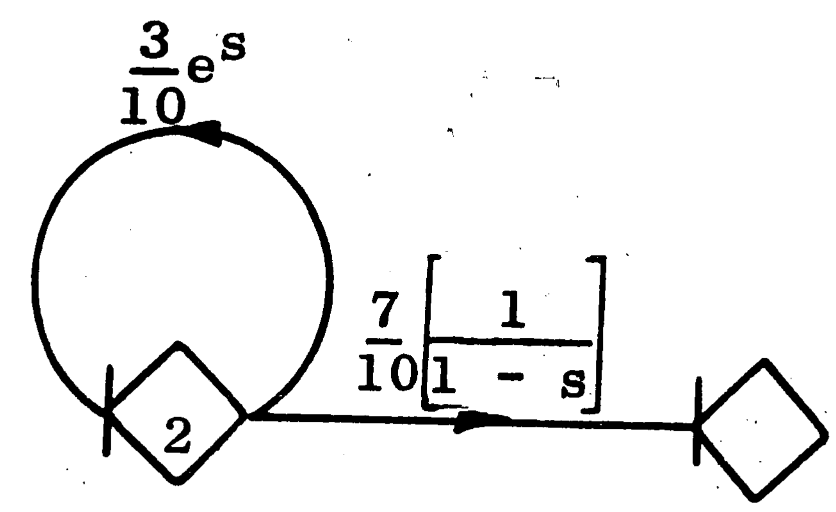
$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_{r} = \left[\frac{d}{ds}M(s)\right]_{s=0} = -\frac{15}{2}$$

$$\therefore q_2^{\,1} = \frac{\mu_r}{\mu_{rt}} = \frac{-\frac{15}{2}}{\frac{5}{2}} = -3$$

ALTERNATIVE 2

Mean Recurrence Time

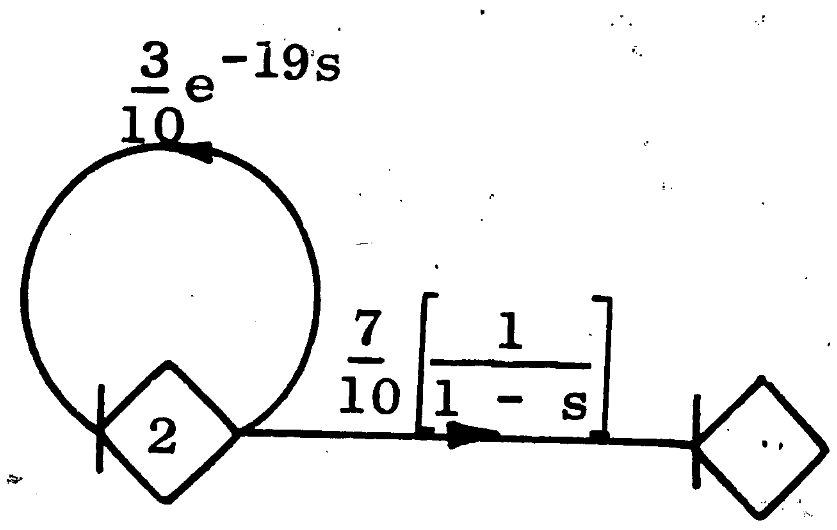


$$W(s) = \frac{\frac{7}{10}\left[\frac{1}{1-s}\right]}{\left[1 - \frac{3}{10}e^{s}\right]}$$

$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_{rt} = \left[\frac{d}{ds}M(s)\right]_{s=0} = \frac{10}{7}$$

Mean Reward



$$W(s) = \frac{\frac{7}{10}\left[\frac{1}{1-s}\right]}{\left[1 - \frac{3}{10}e^{-19s}\right]}$$

$$M(s) = \frac{W(s)}{W(s)\big|_{s=0}} = W(s)$$

$$\mu_{r} = \left[\frac{d}{ds}M(s)\right]_{s=0} = -\frac{50}{7}$$

$$\therefore q_2^{\,2} = \frac{\mu_r}{\mu_{rt}} = \frac{-\frac{50}{7}}{\frac{10}{7}} = -5$$

Figure 14b.. GERT Representation for Calculating Expected Expected Immediate Reward in State 2.

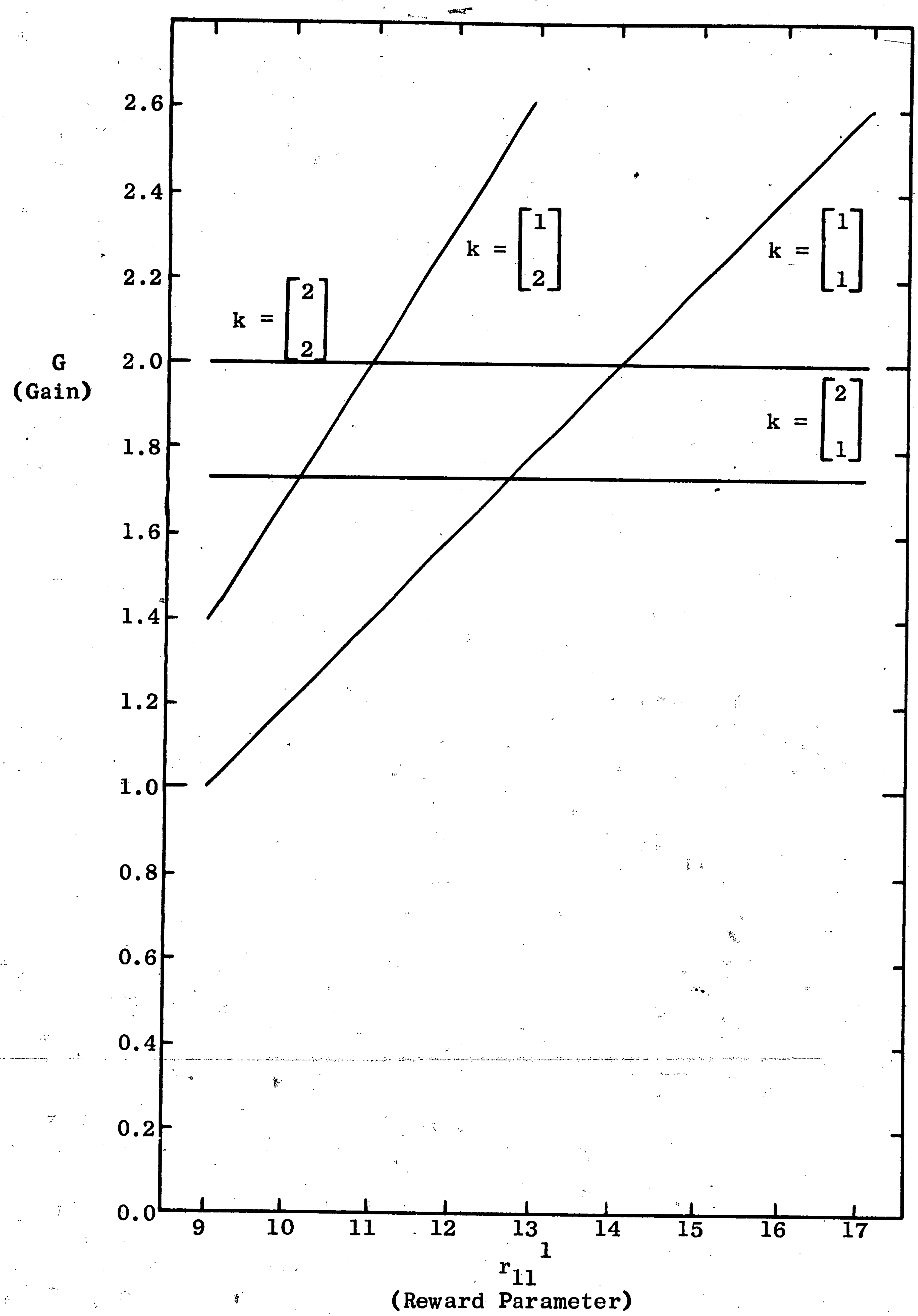
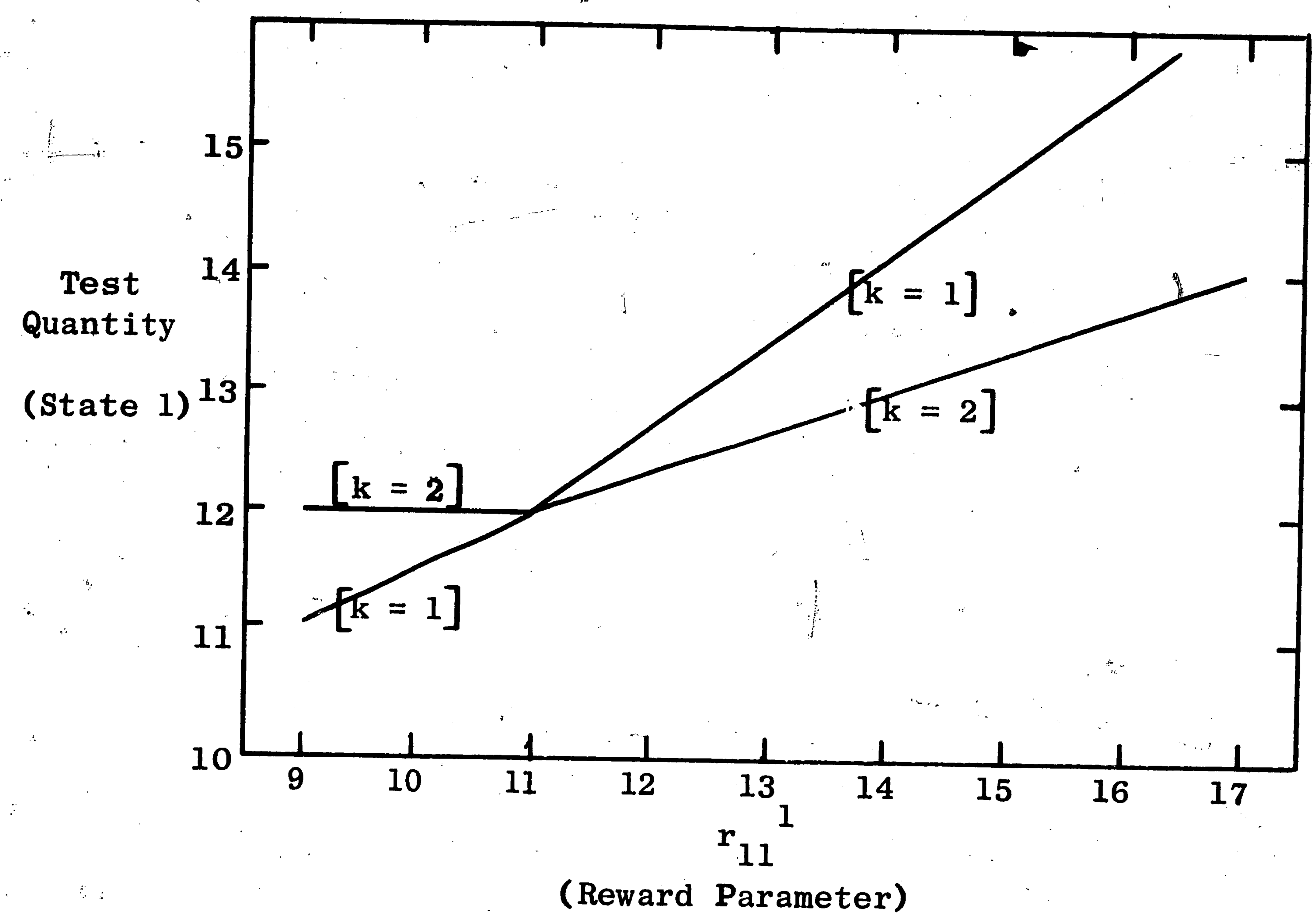


Figure 15. A Graph of the Effects of the Gains for Each of the Four Policies as the Reward Parameter is Varied

$r_{11}^{1} \geq 11$.

Figure 16 is a graph of Howard's test quantities for two states and the two alternatives within each state, as the reward r_{11}^{1} increases from 9 to 17 units. The test-quantities indicate that policy $k = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ does become optimal when $r_{11}^{1} \geq 11$. It follows that the policy iteration technique is sensitive to optimizing problems that are defined by continuous-time transitions.

The means of the exponential distribution chosen to represent the continuous-time transitions in the modified toymaker problem were made equal to the discrete, constant-time transitions in the original toymaker problem. The result is that the rates of return for the four different policies in the continuous-time problem are the same as those calculated for the discrete-time problem. This is as expected, since the means of continuous-time distributions to be combined are additive in nature as are constant-times. The mean of the resultant distribution is therefore the same as that obtained by combining constant-time parameters. It is concluded then that the policy iteration technique will give the intended results for processes that are defined by any transition time distribution as long as the means are known. The means are converted directly into constant-time parameters and the system is solved as if it were a discrete-time process. The point is that the policy iteration technique does not consider the risk (variance) of the combined transition time distribution of the different policies when searching for the optimal policy.



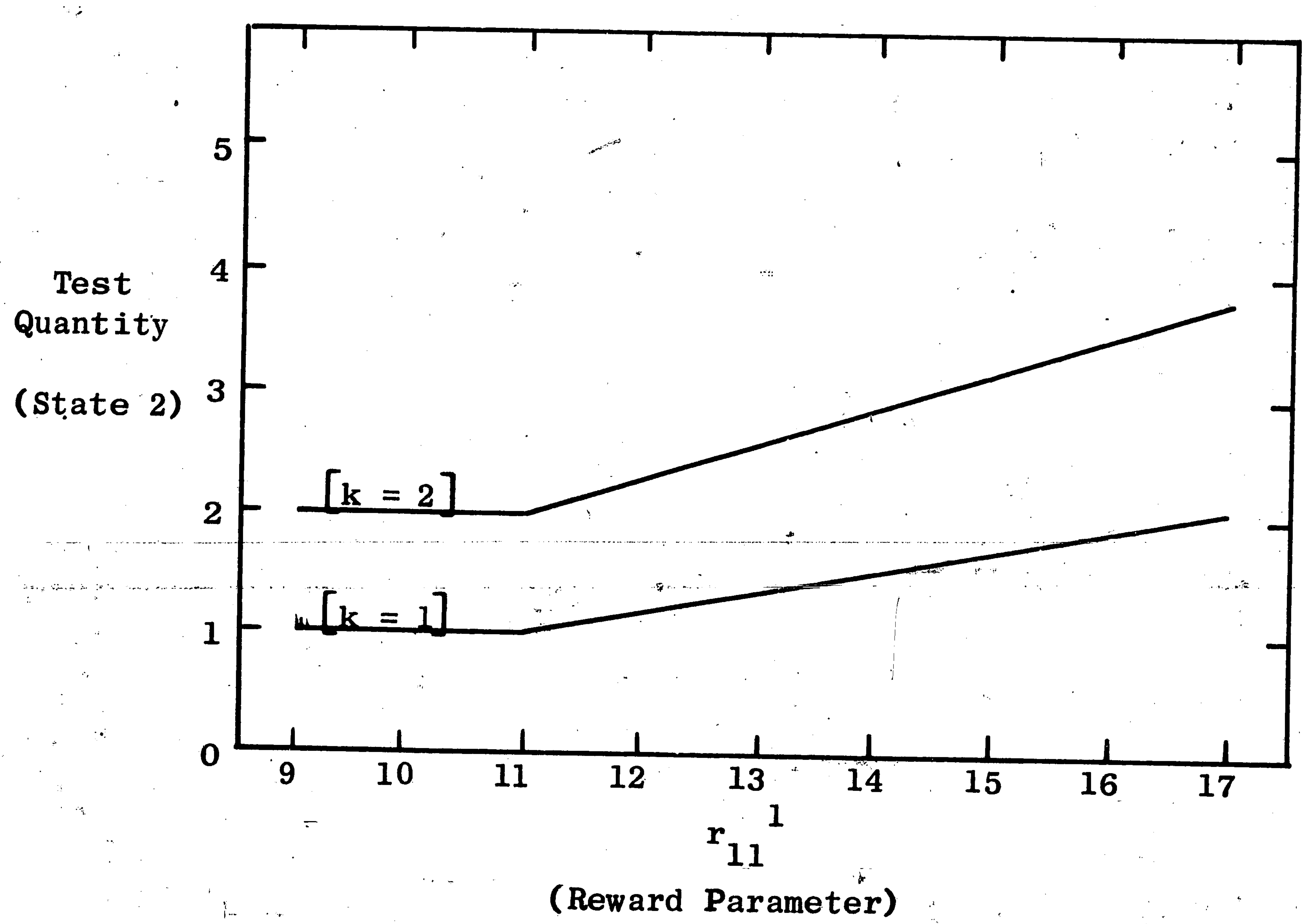


Figure 16. A Graph of the Test Quantities of Howard's Policy Iteration Routine as the Reward Parameter is Varied

Sensitivity Analysis of Other Iterative Techniques

The sensitivity of other iterative techniques may be analyzed in the same manner as the policy iteration routine was analyzed. A gain iteration technique was investigated that closely resembles the policy iteration routine. The difference in the two techniques is in how the test-quantity is calculated. The test quantity used to evaluate the alternative policies at each state is calculated as follows:

$$\sum_{j=1}^{N} p_{ij}^{k} q_{ij}^{k}$$

The results of this test-quantity, as it applies to the modified toymaker problem is plotted in Figure 17 as the reward parameter r_{11}^{1} is again varied from 9 to 17 units. It is noted that this test quantity does not give its intended results since policy $k = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is indicated as optimal where $r_{11}^{1} \leq 13.4$. It is interesting to note, however, that policy $k = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is indicated as the optimal policy by the gain-iteration routine when $r_{11}^{1} \geq 13.4$. This is the same point that policy $k = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ becomes more optimal than policy $k = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

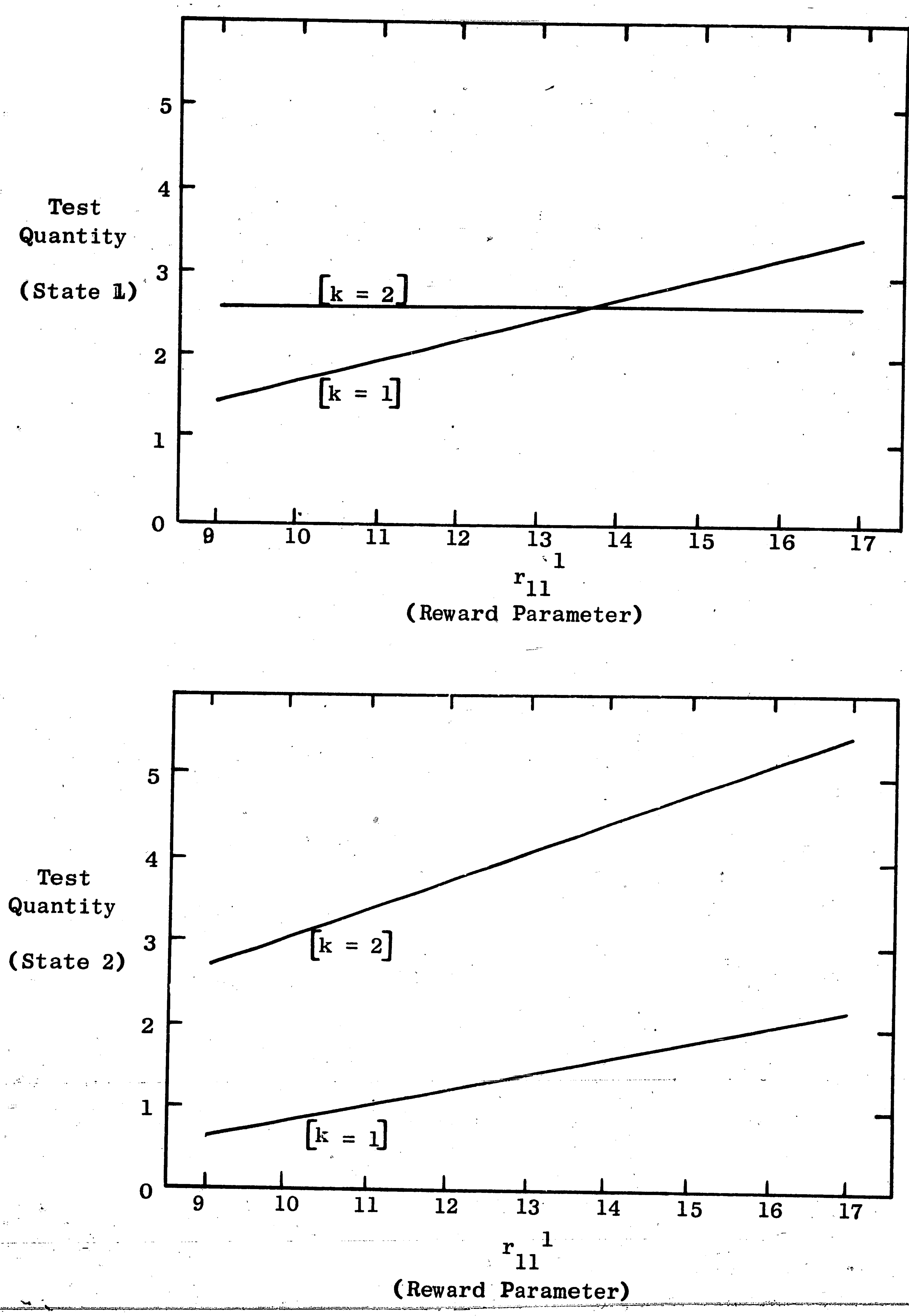


Figure 17. A Graph of the Test Quantities of the Gain-Iteration Routine as the Reward Structure is Varied.

CHAPTER VI

CONCLUSIONS

The purpose of this paper was to extend the usefulness of GERT as a procedure for obtaining optimum results. The approach began by reviewing GERT representations of problems which can be characterized as having Markov transitions.

When a structure of rewards was added to the process, GERT representation proved to be useful for calculating total expected reward and immediate expected rewards. For a system operating under a fixed policy, a knowledge of the total expected reward of the process constitutes a complete understanding of the system.

As alternative policies were made available for the operation of the system, it was shown how GERT can be applied to determine the parameters necessary in Howard's policy iteration technique which locates the optimal policy. It was further found that GERT may be employed directly in continuous time systems as easily as it is used for discrete time systems. This enables the policy iteration routine to be generalized for application in optimizing continuous time systems defined by any transition time distribution, as long as the mean of the distribution is known.

Areas for Future Study

It was pointed out that the policy iteration routine does not consider the variance of the combined distribution of the different policies when searching for the optimal policy. It therefore seems desirable to use GERT in defining a technique that locates an optimal policy that takes into account the variance as well as the mean of combined distributions, and give a desired balance of the risk of the various policies.

GERT has been used exclusively in this paper as an aid in optimizing processes of long duration. Further work is implied in utilizing GERT to analyze processes of short duration; ie, processes with a finite number of state transitions.

Finally, there has been a great deal of interest in the area of maximizing the amount of information that can be made available from a system. Application in this area is possible after making the following considerations:

(1) Is the system Markovian in nature?

(2) Are data necessary to describe the alternatives of the system available?

If the answers to these questions are affirmative, then a possible application has been discovered.

BIBLIOGRAPHY

1. Arrow, K. J., S. Karlin and H. Scarf, Studies in Applied Probability and Management Science, Stanford: Stanford University Press, 1962.

2. Barlow, R., Applications of Semi-Markov Processes to Counter and Reliability Problems, Technical Report No. 57, Stanford University, 1960.

3. Bharucha-Reid, Elements of the Theory of Markov Processes and Their Applications, New York: McGraw-Hill, 1960.

4. de Cani, John S., "A Dynamic Programming Algorithm for Embedded Markov Chains when the Planning Horizon is at Infinity," Management Science, Vol. 10, No. 4, July, 1964, pp. 716-733.

5. Derman, C. and M. Klein, "Some Remarks on Finite Horizon Markovian Decision Models," Operations Research, Vol. 13, No. 2, 1965, pp. 272-278.

6. Feller, W., An Introduction to Probability Theory and its Applications, New York: John Wiley and Sons, Inc., 1950.

7. Howard, R. A., Dynamic Programming and Markov Processes, New York: John Wiley and Sons, Inc, 1960.

8. ____________, "System Analysis of Semi Markov Processes," I.E.E.E. Transactions on Military Electronics, April, 1964, pp. 114-124.

9. Jewell, W. S., "Markov Renewal Programming I: Formulation of Finite Return Models," Operations Research, Vol. 11, No. 6, 1963, pp. 938-948.

10. ____________, "Markov Renewal Programming II: Infinite Return Models, Example," Operations Research, Vol. 11, No. 6, 1963, pp. 949-971.

11. Kemeny, J. G., and J. L. Snell, Finite Markov Chains, Production: Van Nostrand Co., Inc., 1960.

12. Levy, P., "Procesus Semi Markoviens," Proceedings of the International Congress of Mathematicians, (Amsterdam), 1954, pp. 416-426.

13. Lorens, C. S., Flowgraphs for the Modeling and Analysis of Linear Systems, New York: McGraw-Hill Co., 1964.

14. Mason, S. J., and H. J. Zimmerman, Electronic Circuits, Signals and Systems, New York: John Wiley and Sons, Inc., 1960.

15. Morse, P. M., Markov Processes, Chapter 3 of Notes on Operations Research 1959, Cambridge: The Technology Press, M.I.T., 1959.

16. Pargen, E., Stochastic Processes, San Francisco: Holden-Day, Inc., 1962.

17. Pritsker, A. A. B. and W. W. Happ, "GERT: Generalized Evaluation and Review Technique: Development," Submitted to The Journal of Industrial Engineering, 1965.

18. ________________ and G. E. Whitehouse, "GERT: Generalized Evaluation and Review Technique: Applications," Submitted to The Journal of Industrial Engineering, 1965.

19. Pyke, R., "Markov Renewal Processes: Definitions and Preliminary Properties," Annals of Mathematical Statistics, Vol. 32, 1961, pp. 1231-1242.

20. ________, "Markov Renewal Processes with Finitely Many States," Annals of Mathematical Statistics, Vol. 32, 1961, pp. 1243-1259.

21. Smith, W. L., "Regenerative Stochastic Processes," Journal of the Royal Statistical Society Series B, Vol. 20, 1958, pp. 243-302.

22. Takacs, L., Stochastic Processes, Problems and Solutions, New York: John Wiley and Sons, Inc., 1960.

23. Whitehouse, G. E., "Extensions, New Developments, and Applications of GERT," Submitted to Arizona State University, 1965.

VITA

Personal History:

Name:	Gary Raleigh Werberger
Birth Place:	Seattle Washington
Birth Date:	April 17, 1937
Parents:	William F. and Eva G. Werberger
Wife:	Elinor Ann Werberger

Institutions Attended:

Lincoln High School Seattle Washington	Graduated	1955
University of Washington Bachelor of Science Degree Electrical Engineering	Graduated	1959
New York University Business Administration	Attended	1963
Lehigh University Candidate for Master of Science Degree Industrial Engineering		1966

Experience:

Electrical Engineering Officer in Charge of Navigational Aids Ashore, United States Coast Guard, San Francisco, Calif.
Development Engineer, Western Electric Co., New York, N. Y.
Research Engineer, Western Electric Co., Princeton N. J.